Mancero, Gabriela (2010) Detection of changes through visual alerts and comparisons using a multi-layered display. PhD thesis, Middlesex University. [Thesis]

DETECTION OF CHANGES THROUGH VISUAL ALERTS AND COMPARISONS USING A MULTI-LAYERED DISPLAY

Maria Gabriela Mancero

A thesis submitted to Middlesex University in partial fulfillment of the requirements for the degree of Doctor of Philosophy

School of Engineering and Information Sciences

Middlesex University

May 2010

MIDDLESEX UNIVERSITY **FORM RD(Decl)**

Candidate's Declaration Form

Name of candidate: Maria Gabriela Mancero A.

Thesis Title: Detection of changes through visual alerts and comparisons using a Multi-Layered Display

Degree for which thesis is submitted: Doctor of Philosophy (Engineering and Information Sciences)

1 Statement of associated studies undertaken in connection with the programme of research (Regulation G3.1 refers)

While a registered student of the University, I attended the following courses / workshops / conferences:

ECCE 2007 (Conference and Doctoral Consortium) - August 28-31, London, UK

10^{th} Human-Centred Technology Postgraduate Workshop - December 6, Brighton, UK

HFES 2008 - September 22-26, New York, USA

ECCE 2009 - November 30 – October 2, Helsinki, Finland

OzCHI 2009 - November 23-27, Melbourne, Australia

2 Concurrent registration for two or more academic awards

I declare that while registered as a candidate for the University's research degree, I have not been a registered candidate or an enrolled student for an award of another university, academic or professional institution.

3 Material submitted for another award

I declare that no material contained in the thesis has been used in any other submission for an academic award

Signature of candidate .... Date ....

Abstract

The Multi-Layered Displays (MLD) comprise two LCD screens mounted one in front of the other, allowing the presentation of information on both screens. This physical separation produces depth without requiring glasses.

This research evaluated the utility of the MLD for change detection tasks, particularly in operational environments. *Change Blindness* refers to the failure to detect changes when the change happens during a visual disruption. The literature equates these visual disruptions with the types of interruptions that occur regularly in work situations. Change blindness is more likely to occur when operators monitor dynamic situations spread over several screens, when there are popup messages, and when there are frequent interruptions which are likely to block the visual transients that signal a change.

Four laboratory experiments were conducted to evaluate the utility of the MLD for change detection tasks. The results from the experiments revealed that, when depth is used as a visual cue, the depth of the MLD has a different effect on the detection of expected changes and unexpected events. When the depth of the MLD is used as a comparison tool, the detection of differences is limited to translation differences in simple stimuli with a white background. These results call into question previous claims made for the MLD regarding operational change detection.

In addition, observations and interviews were used to explore whether change blindness occurred in an operational command room. The results suggested that operators develop strategies to recover from interruptions and multitasking. These results call into doubt the wisdom of applying change detection theories to real world operational settings. More importantly, the research serves as a reminder that cognitive limitations found in the laboratory are not always found in real world environments.

Acknowledgements

This research involved many years of effort and would not have been possible without the assistance of great many people. I want to express my gratitude to my supervisors, friends, family and colleagues whose support kept me going through.

I particularly want to thank Martin Loomes, Paul Cairns and Richard Comley. Martin's attention to detail made my final work much better than it otherwise would have been. His experience allowed me to expand my knowledge base and his ability to foresee problems has helped me improved my analytical skills. His assistance has made this thesis possible and I truly believe that his influence has helped me to become a better researcher. My heartfelt appreciation goes to Paul Cairns and Richard Comley who were the source of an excellent discussion.

I would like to thank the European Organization of Airspace Research and Development (EOARD), the United States Air Force Research Laboratory (AFRL), Award Grant No. FA8655-06-1-308, and Middlesex University for sponsoring this project.

Special thanks to the British Transport Police staff, in particular to Paul Garrett, Nigel Griffin and Douglas Robbins who provided me with invaluable assistance and experience in the control room.

I would like to acknowledge Bali Rooprai, Robert O'Shea and all those blind reviewers for their valuable input.

Finally I would like to thank Stephen Gaukrodger for encouraging me to pursue promising topics of interest, for his invaluable comments, his help with programming, and for putting up with me not only while writing the thesis, but always. I truly believe that he played a pivotal role during the development of this research and I would not have made it without him.

Thank you all for your help.

Publications related to this PhD.

2009

Mancero, G., Wong, B. L. W., & Loomes, M. (2009a, November 23-27). *Radio Dispatcher's Interruption Recovery Strategies.* Paper presented at the OzCHI 2009, Melbourne,Australia.

Mancero, G., Wong, W., & Loomes, M. (2009b, November 30 - October 2). *Change Blindness and Situation Awareness in a Police C2 Environment.* Paper presented at the European Conference on Cognitive Ergonomics (ECCE), Helsinki, Finland.

Dunser, A., & **Mancero, G. (2009).** The use of Depth in Change Detection and Multiple Object Tracking. *The Ergonomics Open Journal, 2*, 142-149.

2008

Mancero, G., & Wong, B. L. W. (2008). *An Evaluation of Perceptual Depth to Enhance Change Detection.* Paper presented at the Human Factors and Ergonomics Society 52nd Annual Meeting

Mancero, G., Wong, B. L. W., & Amaldi, P. (2008). *The Utility of Depth in Multi-Layered Displays for Supporting Change Detection* (No. 2008-4-001). London: Middlesex University and European Office of Aerospace Research and Development.

2007

Mancero, G., Wong, B. L. W., & Amaldi, P. (2007a). *Looking but Not Seeing: Does Perceptual Depth Reduce Change Blindness?* (No. Technical Report 2007-3-001). London: Middlesex University and European Office of Aerospace Research and Development.

Mancero, G., Wong, B. L. W., & Amaldi, P. (2007b, 28-31 August). *Looking but Not Seeing: Implications for HCI.* Paper presented at the European Conference on Cognitive Ergonomics (ECCE), Covent Garden, London.

Table of Contents

List of Figures

List of Tables

Chapter 1

INTRODUCTION

In their early stages, innovative technologies may hint at new ways to work, play and live. Unfortunately, new technologies often require many years of research and development before they find a niche.

The Multi-Layered Display (MLD) is one such technology. It comprises two LCD screens mounted one in front of the other. It allows the presentation of information on both screens by making those images on the front layer translucent. The overlapping screens produce true depth without the need for glasses (Figure 1-1).

Figure 1-1. The MLD, a side-view of the MLD presenting data in each layer and a schematic representation of the MLD's layout.

Developed in 1999 by PureDepth, the MLD was a speculative technological innovation rather than a solution for an existing problem. It was not until 2007 that PureDepth launched its first commercial product, a multi-layered casino slot machine (2007) which was introduced to the Japanese market in November 2008 (BusinessWire, 2009).

Between 1999 and 2008, PureDepth contracted a variety of research organisations to search for functional applications for the MLD. These investigations identified some potential uses of the MLD, including the ability to segregate information in different depth planes, to present details on the front screen as the focus of attention and keep the context on the rear screen, or to improve the users' search performance (Wong, Mansour *et al.*, 2005). However, results from empirical testing reported in the literature, generally using visual search or multiple object tracking paradigms, are inconclusive (Aboelsaadat *et al.*, 2004; Bishop, 2005; Bolia *et al.*, 2004; Carr *et al.*, 2006; Dunser *et al.*, 2008; Hayes *et al.*, 2006; Masoodian *et al.*, 2004).

Furthermore, PureDepth's internal reports suggested that the MLD may be used to "solve" the change blindness problem (PureDepth, 2007; Singh, 2005). They stated that the MLD's depth could enhance change detection by segregating information by layer to highlight important information or allowing comparison of images (PureDepth, 2007; Singh, 2005).

This PhD project was funded through a grant[1] from the European Organization of Airspace Research and Development (EOARD) and the United States Air Force Research Laboratory (AFRL) to evaluate the utility of the MLD for change detection tasks assuming that change blindness is a problem in operational environments. This last premise was justified by previous research suggesting that change blindness is likely to occur when operators are frequently interrupted, monitor several screens and perform various tasks simultaneously (Di Vita *et al.*, 2004; Durlach, 2004b; Podczerwinski *et al.*, 2002; Smallman *et al.*, 2003; St. John *et al.*, 2005; St. John *et al.*, 2007; Yeh *et al.*, 2000). Studies on change blindness in operational environments have used part-task simulations to measure the magnitude of change blindness. Hence, an implicit part of this project is to test this underlying assumption and determine if change blindness is likely to occur in operational settings.

[1] Award Grant No. FA8655-06-1-3081

1.1. Change Detection as a possible application

Two cognitive phenomena known as *change blindness* and *inattentional blindness* have shown that we fail to detect changes that seem highly noticeable under normal conditions, like the appearance of a gorilla in the middle of a basketball game (Simons *et al.*, 1999), or the disappearance of an airplane engine in a picture where the airplane is the main focal point (Rensink *et al.*, 1997). One would have thought that the appearance of a gorilla or the disappearance of an airplane's engine would be easy to spot, but these changes may not be detected unless one is paying attention to the specific objects or events that change.

Inattentional Blindness refers to the failure to notice unexpected events that appear within our field of view when our attention is diverted to a different task (Mack *et al.*, 1999). Evidence of inattentional blindness in the real world comes primarily from studies that have investigated the effect of using mobile phones while driving (Herslund *et al.*, 2003; Mack, 2007; Scholl *et al.*, 2003). Results from these studies have shown that using a mobile phone while driving dramatically increases the chances of not detecting potential changes such as an incoming bicycle.

Change Blindness refers to the failure to detect a change that occurs within our field of view and happens during a visual disruption (Rensink *et al.*, 1997). If there is no visual disruption, change blindness could also occur if the change happens at a very slow rate (Simons *et al.*, 2000). Evidence of change blindness in operational environments comes from laboratory experiments using part-task simulations of army, military and flight monitoring systems in which participants have shown poor detection rates (Di Vita *et al.*, 2004; Durlach, 2004a, 2004b; Muthard *et al.*, 2003; Podczerwinski *et al.*, 2002; Pringle *et al.*, 2001; Smallman *et al.*, 2003; St. John *et al.*, 2005). Although few of these studies have used domain experts, their results have been taken as a confirmation that the phenomenon occurs in command and control environments.

The literature suggests that operators in command and control rooms are usually loaded with visual search, voice communications and situation assessment tasks (Di Vita *et al.*, 2004; Durlach, 2004b). According to Durlach (2004b), in such environments, multitasking can slow change detection. Di Vita et al. (2004) stated that operators who work with several monitors might miss time-critical information because the process of shifting their attention from one display to another creates an opportunity for changes to occur on unattended screens. It has also been suggested that in tasks where users monitor dynamic situations, from air traffic management to civil emergency response coordination, interruptions disrupt users' situation awareness and cause them to miss important changes (Podczerwinski *et al.*, 2002; Smallman *et al.*, 2003; St. John *et al.*, 2005; St. John *et al.*, 2007; Yeh *et al.*, 2000). According to Di Vita (2004), this problem can be resolved if the interface specifically draws attention to the changed information when the operator resumes viewing a previously unattended display.

1.2. Implications for HCI

Many man-made systems and processes rely heavily on visual displays to convey information, so systematic failures in the detection of changing information have implications for human-computer interface design (Varakin *et al.*, 2004). Researchers have suggested that interface designers should take account of change blindness using efficient visualizations and algorithms that present information and alert the operators to changes that occurred while their attention was diverted (Di Vita *et al.*, 2004; Durlach, 2004b; Smallman *et al.*, 2003). The laboratory evidence for visual lapses and the apparent need for new ways to draw operators' attention to important visual changes present a challenge and an opportunity for the MLD. The use of depth could be an alternative to traditional visual alerts.

1.3. Key Concepts

The funding for this research focused attention on improving human change detection performance. It frames the research on the cognitive aspects of change detection. This research project does not investigate the neurological aspects of how the brain reacts to change detection tasks (Beck *et al.*, 2001; Curran *et al.*, 2009; Pessoa *et al.*, 2004) or computational algorithms to improve digital detection of changes, as in remote sensing data (Bruzzone *et al.*, 1999; Lu *et al.*, 2004).

This research project focuses on the evaluation of the MLD as a tool to enhance visual change detection performance. For the purpose of this research project, visual changes are limited to a modification of colour, orientation, displacement or existence of a stimulus.

Cognitive psychologists have differentiated between change perception and difference perception. Change perception refers to the detection of an ongoing transformation of a structured object. It is a variation referenced to a structure (Rensink, 2002) and refers not only to the visual processes involve in noticing a change (detection), but also its identification and localization (Rensink, 2002). Difference detection refers to the inferential comparison of the current stimulus with traces in long-term memories (Rensink, 2002; Scott-Brown *et al.*, 2000; Simons & Rensink, 2005).

While the detection of changes requires focused attention and the detection of differences requires memory, both concepts were necessary to investigate the utility of the MLD for change detection tasks. In situations where the system alerts the operator about a change, the system needs to capture the operator's attention to the changing stimuli. Thus, theories of change detection became relevant. However, if the computer is not used to mediate detection, then the technology should allow the operator to directly compare the stimuli to detect the changes. Then, studies on difference perception were applicable.

Another important distinction is that between *expected* and *unexpected* changes. An expected change refers to the modification of a stimulus which the user or operator is waiting for. The user does not know what type of change will occur or when it will occur but knows that a change will happen. An unexpected change is the appearance of a stimulus that is not anticipated by the user or operator. The user does not know that something will come up on the screen.

A third distinction is that between the detection of *dynamic* and *completed* changes. The detection of dynamic change refers to the perception of the transformation itself, while the detection of completed change refers to the perception that the structure changed at some point in the past (Rensink, 2002).

1.4. Problem Statement

The grant received for this project established that one possible application for the MLD in command and control environments is to enhance the detection of visual changes. The previous sections have outlined the research domain and the motivation behind this work. From these, the primary research problem can be stated as follows:

> Given that (a) there is evidence for change blindness in the laboratory and that (b) operators in command and control rooms might be blind to changes; is the depth of the MLD a useful tool for change detection tasks that could potentially be applied to command and control rooms?

1.5. Overview of the Methodology

To address the stated problem, two fundamental issues needed to be evaluated: The first one was to evaluate the utility of the technology for change detection tasks. The second addressed the cognitive phenomena in an operational environment.

This research used both qualitative and quantitative methods to address these issues. Quantitative methods were used to address the utility of the MLD for change detection in laboratory settings. Qualitative methods were used to search for change blindness in an operational environment. The testing of the two issues was treated as a pair of separate research questions. The question of the MLD's utility was best answered with statistical analysis, while that of its necessity was more amenable to inductive analysis.

The evaluation of the utility of the MLD raised two questions regarding the effectiveness of depth to guide attention to crucial information and to different foci to allow comparison:

(a) ***Can the depth of the MLD be used as an alerting tool?***

(b) ***Can the depth of the MLD be used as a comparison tool?***

The exploration of the cognitive phenomena in an operational environment also raised two questions regarding the vulnerability of operators to change blindness when multitasking during demanding times or after an interruption:

(c) ***Do operators miss changes when multitasking?***

(d) ***Do operators miss changes after an interruption?***

It would have been preferable to conduct the field study first to address *(c)* and *(d)* but there were deadlines stipulated in the contract with the sponsors which were related to *(a)* and *(b)*. Additionally, obtaining permission to access a control room took almost a year.

The literature on empirical testing of the MLD had paid little attention to change detection tasks. To determine whether the depth of the MLD has an effect on change detection, four laboratory experiments were conducted. The first two experiments were concerned with situations where the computer highlights the changes to the operator. They address question *a*. Experiment

1 evaluated the effect of depth to highlight expected changes. Experiment 2 investigated the effect of depth to highlight unexpected events. An analysis of the effect of depth in change detection tasks was presented at the 52nd Annual Meeting of the Human Factors and Ergonomic Society (Mancero & Wong, 2008). A report on the utility of the MLD was presented to EOARD - the European Organization of Airspace Research and Development (Mancero, Wong *et al.*, 2008). The overall results of both experiments are described in detail in Chapter 5.

The last two experiments hypothesised that, instead of the computer mediating the detection of changes, the technology should allow the observer to directly detect differences between two sets of stimuli. Experiments 3 and 4 explored whether the depth of the MLD could be used as a comparison tool using simple and complex images. They comprised the detection of completed changes. These experiments address question *b*. The overall results are described in Chapter 6.

Access was gained to a command and control environment that had the characteristics described in the literature: The British Transport Police (BTP) Force Control Room in London, United Kingdom. Once access was obtained to a control room, the next step was to investigate the phenomena of change and inattentional blindness in the operational environment. It was necessary to explore whether operators in the command and control room were 'blind' to changes, and if so, determine if the MLD could be used in this operational environment. The field study involved extensive interviews, observations and review of internal documents and performance indicators. The results address questions *c* and *d*. An analysis of the difficulties faced by BTP radio dispatchers especially when multitasking was presented at the European Conference on Cognitive Ergonomics (ECCE) 2009 (Mancero *et al.*, 2009b). The description of interruption recovery strategies developed by the BTP radio dispatchers was published at the proceedings of the Australian Computer-Human Interaction (OzCHI) 2009 conference (Mancero *et al.*, 2009a). The overall results of the field study are reported in Chapter 7.

In summary, this research used both qualitative and quantitative methods to address the problem statement. The quantitative methods were used to address the utility of the MLD for change detection in laboratory settings, while qualitative methods were used to explore change blindness in an operational environment. Table 1-1 provides a summary of the aims and methods used for this research:

Table 1-1. Summary of studies carried out, aims and methods

Experiment/ Study	Aim	Method	Participants
Experiment 1	Evaluate whether the depth of the MLD has an effect on detection if used to highlight expected changes	Controlled Experiment	22
Experiment 2	Evaluate whether the depth of the MLD has an effect on detection if used to highlight unexpected events	Controlled Experiment	60
Experiment 3	Evaluate whether the MLD's depth has an effect on the detection of differences with simple stimuli	Controlled Experiment	24
Experiment 4	Evaluate whether the MLD's depth has an effect on the detection of differences with complex stimuli	Controlled Experiment	24
Field Study Part 1	Determine whether operators miss changes when multitasking	Observations Critical Decision Method Review of internal reports	7
Field Study Part 2	Determine whether operators miss changes after an interruption	Observations Critical Decision Method Review of internal reports	5

1.6. Key Challenges

A major challenge was to develop a methodology that would adapt to the situation where many of the factors usually determined using a typical user-centred design process were predetermined by the manufacturers of the MLD and the sponsors of the project. A typical user-centred design process consists of four phases: Analysis, Design, Implementation and Deployment (UPA, 2009; Vredenburg *et al.*, 2002). During the analysis phase, the first step is to meet with the stakeholders and determine the requirements. However, the nature of the funding meant that the requirements, improve the detection of visual changes, had already been established, and that it fell to the researcher to find stakeholders in an operational environment that would benefit from this improvement.

Having found and gained access to a command and control room that fitted the requirements described in the literature, a not insubstantial task, led to the next problem: to choose an investigative method that would support the detection and systematic measurement of change blindness in an operational context. The researcher conducted extensive observations and interviews but not electronic equipment was permitted.

Once work in the operational environment began, the next challenge was to define operational change. However, any definition is hard to generalize because the "importance" of a change is context-dependent – changes that are important in one situation might not impact upon another –. Visual changes that are important for a radio dispatcher in control room A might not be important to a radio dispatcher in control room B.

1.7. Structure of the Thesis

Chapters 2, 3 and 4 focus on different aspects of the problem domain: Chapter 2 introduces the technology, its capabilities, limitations and the gaps in the empirical research found in the literature. Chapter 3 provides an overall analysis of human depth perception. This analysis is essential for any discussion on display technologies that produce a three dimensional effect because the mechanical aspects of the technology and those that underlie depth perception combine to define the limits of the overall system. Chapter 4 introduces the problem domain that the system will be applied to. This chapter reviews the literature on change blindness and visual attention to identify visual cues that capture attention and the requirements for change blindness to occur in an operational environment.

Chapters 5 and 6 describe basic research conducted on the MLD. Chapter 5 describes two laboratory experiments that evaluate whether the depth of the MLD use as a visual cue has an effect on the detection of expected changes and unexpected events. Chapter 6 explores the use of depth to allow comparison of images. It evaluates the effect of the MLD's depth on the detection of differences between simple images like basic shapes and more complex ones like photographs.

Chapter 7 explores the dynamics of the change blindness phenomenon in a command and control room. It describes a field study conducted at the British Transport Police (BTP) and the strategies that BTP operators developed to recover from interruptions and multitasking.

Chapter 8 provides a series of conclusions and lessons learnt during the execution of this research project. It discusses the research limitations and the possibilities for further research.

1.8. Summary

The MLD is a display technology that comprises two LCD screens one in front of the other. This architecture allows to present information on both screens that are separated with real depth. The developers of the MLD claimed that this display could solve the "change blindness" problem in operational environments due to its capability to segregate information on different layers, allowing to highlight changes or to compare images. This project received a grant to evaluate the utility the MLD for change detection tasks assuming that it could potentially be used to enhance operational change detection.

Change blindness is a widely studied cognitive phenomenon that refers to the failure of noticing a change. People can fail to detect a change if the change happens during a visual disruption. Without a disruption, change blindness could also occur if the change happens gradually. This research aims to determine whether the depth of the MLD is a useful tool for change detection and could have any potential use in a command and control environment.

Chapter 2

THE MULTI-LAYERED DISPLAY

The MLD is a novel display that provides a true three dimensional effect to the user. To make the most of any display technology, it is important to understand the characteristics of the technology and the nature of human vision to define the limits of the overall system.

This chapter introduces the MLD, compares it to other autostereoscopic displays and discusses previous research on the MLD.

2.1. The MLD: Technical Aspects

PureDepth's Multi-Layered Display devices comprise two physically distinct layers of LCD panels sharing a common back-light source. The two LCD layers are separated by a clear Perspex layer with thickness ranging between 1 and 15 mm depending on the model. The two LCDs run from separate VGA cables, which may connect to a single video card. Since the images are at different distances from the viewer, the viewer sees actual depth without the need for special glasses. Figure 2-1 shows a sequence of images as the observer's head is moving where a circle is on the front layer and a cross is on the back layer.

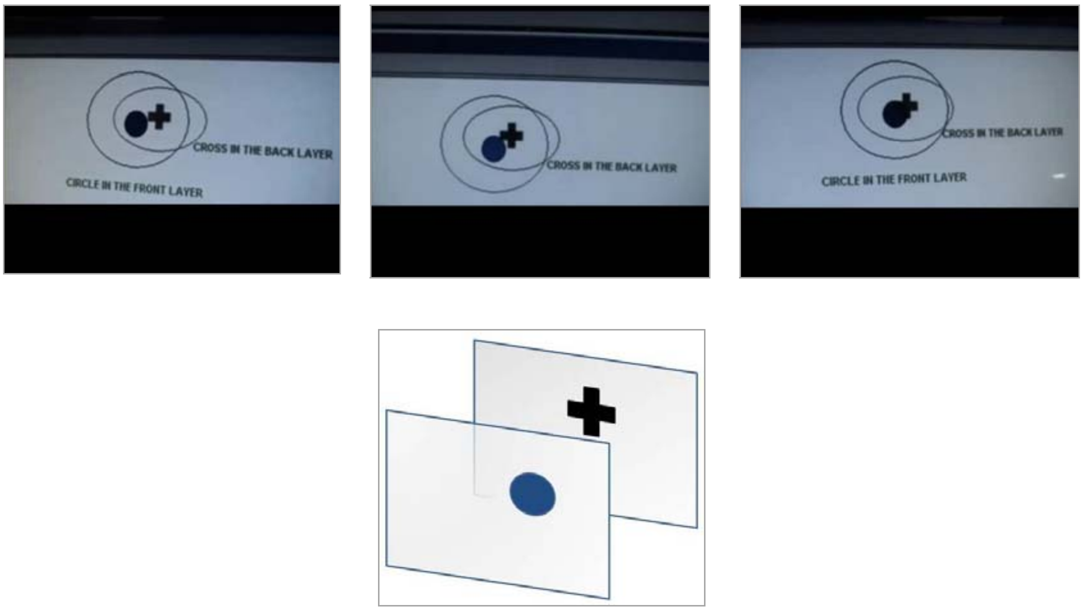


Figure 2-1: Sequence of images of the MLD as head moves and schematic representation

The physical architecture of the MLD produces three main technical problems: moiré interference, loss of a significant proportion of light, and changes in colour and contrast ratio.

Moiré interference occurs whenever two regular patterns of slightly different spatial frequency overlap (Bell *et al.*, 2007) and are slightly misaligned (Figure 2-2). This is a problem for the MLD, because of limits in manufacturing precision. To counteract the moiré interference, a diffuser element was added between the two screens. Although it minimizes the moiré interference, it leads to degradation of the rear LCD image quality by slightly blurring the images (Bell *et al.*, 2007).

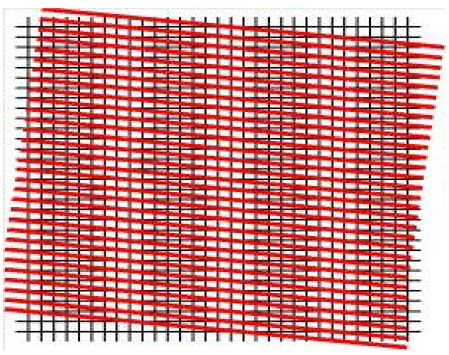


Figure 2-2. Moiré Interference. Source: http://tinyurl.com/yf38n9c

In addition, the presence of two screens leads to the loss of a significant proportion of light. This loss is between 95 and 98%. As a result, multi-layered displays tend to have a luminance of 150 cd/m^2 as compared to a regular display with a luminance of 300-400 cd/m^2 (Yahiro *et al.*, 2006). To solve the luminance drop, the backlight of a multi-layered display is significantly brighter than that in an ordinary display to provide an equivalent luminance to a traditional single-layer display (Bell *et al.*, 2006; Bell *et al.*, 2008). This of course has a negative effect on the efficiency of the unit, making the multi-layered display technology unsuitable in situations where power consumption is limited by use of a battery, fuel cell or other storage device.

Stacking two screens also changes the colour gamut and the contrast ratio compared to a single-layered display (Bell *et al.*, 2008). The colour model of the MLD is both additive and multiplicative. In the additive colour model, red, green, and blue (RGB) are the primary colours, and mixing them together creates white (Chadwick, 1999). With multiplicative blending models, the resulting pixel values are the product of the foreground and background pixels (Bunks, 2000). While each layer of the MLD independently uses the additive model, the colours between layers are effectively multiplied (Bishop, 2005). Colours in the front layer become translucent, so white in the front layer is completely transparent. Other colours become intensified if shown in both layers.

The quality of the images could be improved if alternatives to LCDs are used. Other types of screens such as optically-controlled birefringence (OCB) and organic light emitting diodes (OLED) improve the refraction of light but are more expensive than traditional LCDs. OCB displays divide a ray of light into two. Each ray is refracted at a different angle and polarized at a right angle to the other which improves the contrast ratio and reduces crosstalk[2]. On the other hand, OLED displays have a significant advantage over traditional LCDs and OCBs. Because they do not require a backlight to function, they use less power, they can display deep black levels, and can be much thinner and lighter than an LCD panel (PureDepth, 2007).

2.2. Using the MLD

Most modern operating systems support multiple monitors so the MLD does not require special drivers. It is recommended to set up the two displays as if they were side by side like an extended desktop, so data located on the left-hand side is presented on the front layer of the MLD, while the data located on the right-hand side is presented on the rear layer. This set up maps the two layers of 1280×1024 into a single virtual desktop with a resolution of 2560×1024 although the resolution of each display could be changed independently (Figure 2-3).

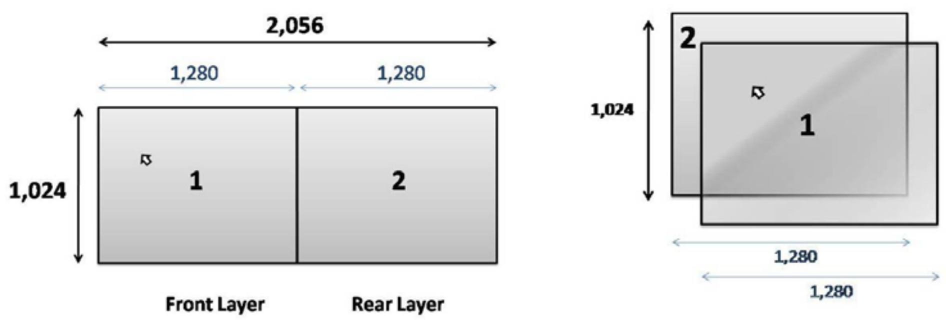


Figure 2-3. MLD Screens Set-up and dimensions

The MLD comes with a utility that allows the user to move the cursor from the front to the rear layer and vice versa by clicking the middle mouse button. Provided that the displays are set as if they were side by side, when the

[2] Crosstalk refers to a signal transmitted on one circuit or channel of a transmission system that creates an undesired effect in another circuit or channel. In 3D displays, it appears to the observer as a ghost image. Hence, it is also known as *ghosting*.

mouse pointer is on the front layer, moving it to the far right will move the pointer to the rear layer until it reaches the end of the screen. Then, the mouse pointer will have to be moved to the far left to get it back to the front layer, and it will stop when it reaches the left end of the front screen.

2.3. The MLD and other 3D Display Technologies

Three-dimensional displays can be categorized by the technique used to channel the left and the right images to the appropriate eye: some devices, known as stereoscopic displays, require optical devices close to the observer's eye, while autostereoscopic displays rely solely on characteristics of the display dispensing the need for user eyewear.

Stereo-based 3D technologies use stereoscopy to create an illusion of depth. Stereoscopic systems direct different images to each eye through angular or polarization multiplexing. They require users to wear a device, such as polarised glasses in combination with a method of polarising the two views; shutter glasses working in synchronisation with a view switching display; anaglyph glasses analysing different colour channels to obtain images; and head-mounted displays (HMD) utilising motion parallax cues for creating an illusion of depth (Holliman, 2006; Naikar, 1998).

Some of the limitations of stereoscopic displays are that the viewer has to wear glasses and the system usually generates raster lines that degrade the vertical resolution of the display (Wickens *et al.*, 1989). For stereoscopic direct view LCD based displays, the main drawback is the crosstalk (see footnote 2) created because of the parallax between the display pixels and the micro-polariser mounted over the LCD (Holliman, 2006). Despite the drawbacks, studies on neurosurgical visualizations have shown that participants achieved best performance when using polarised glasses compared to HMD or multiview lenticular displays (Cooperstock *et al.*, 2009).

On the other hand, autostereoscopic displays are those that do not require the observer to wear any device, sending separate images directly to the correct eye (Holliman, 2006). Autostereoscopic 3D displays produce an optical output that creates a region in space of at least two viewing windows (one for each eye). If an observer places the right eye in one window, and the left eye in another, each eye sees a different image on the display constituting a stereo pair that is seen without the need to wear glasses.

The MLD could be classified as a discrete parallax display due to the positional disparity that the images in each layer has. Each LCD presents two-dimensional images. The physical separation between the MLD screens produces real depth instead of just creating an illusion of depth.

Other parallax display devices consist of a layer which elements can emit light of varying intensity in different directions. The input consists of two-dimensional projections such as photographic or synthetic images that contain no explicit depth information. Instead, depth is implicitly encoded as positional disparity between different projections (Halle, 1997). There are two main optical elements used to generate the appropriate viewing windows for each eye: parallax barriers and lenticular elements.

The parallax barrier consists of an opaque layer of material with a series of regularly spaced vertical slits (Figure 2-4). Usually a piece of film or other imaging medium is offset some distance behind the parallax barrier. Each slit in the barrier acts as a window onto a stripe of the section of the film. The stripe that is visible depends on the horizontal angle from which the slit is viewed (Halle, 1997). The left and the right images are interlaced in columns on the display and the parallax barrier is positioned so that left and right image pixels are blocked from view except in the region of the left and right viewing window respectively, producing a stereoscopic image (Holliman, 2006). The parallax barrier divides the light so that different patterns reach the viewer's left and right eyes creating the perception of a three-dimensional image.

Parallax barrier displays present some limitations: the resolution of the film limits the maximum number of views that can be displayed and the spacing of the slits determines the maximum spatial resolution of the display. Additionally, the parallax panoramagram displays have thinner columns allowing more views behind each slit. However, their three-dimensional effect is only in the horizontal direction, while, vertically, the images behave as if they were flat photographs (Halle, 1997).

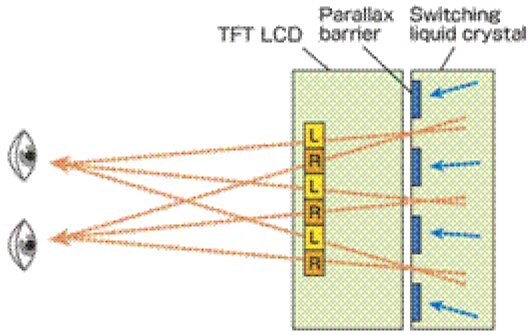


Figure 2-4. Parallax barrier. Source: http://tinyurl.com/ygpow8t

Lenticular elements used in 3D displays are cylindrical, long, narrow lenses, instead of slits, that are arranged vertically with respect to a 2D display such as an LCD (Holliman, 2006). A lenticular lens is an array of magnifying lenses, designed so that when viewed from slightly different angles, different images are magnified (Figure 2-5).

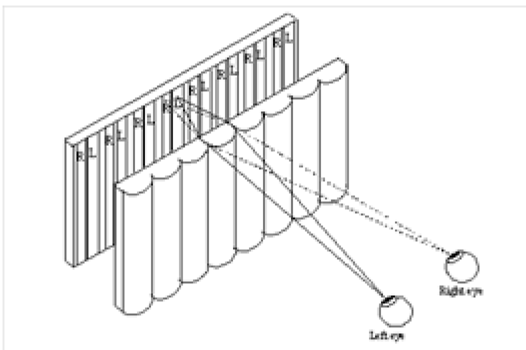


Figure 2-5. Lenticular elements. Source: (Holliman, 2006)

Each lens focuses on the image information located behind it and directs the light in different directions. The entire surface of the lenticular sheet radiates light producing no dark stripes such as those in a parallax barrier (Halle, 1997).

Some limitations of the lenticular displays could be presented due to manufacturing processes: the quality of the lenticular sheet determines the optical aberrations that will be manifested in the final image. The optical power of the lens controls the angle of view through which the final image can be seen, but the higher quality of the lenses, the higher the cost of producing lenticular sheets.

The quality of the image also depends on the spatial resolution of the two-dimensional display that the lenticular element is attached to. The 2D display should be high enough in the horizontal direction to provide both spatial and directional information. Like the parallax panoramagrams, lenticular displays only show horizontal parallax. Optical alignment of the underlying display with the lens sheet is essential to producing distortion-free three-dimensional image.

The position of the lenticular sheet determines the number of viewing windows that a display can produce. Multi-view systems support greater viewing freedom by generating multiple simultaneous viewing windows of which an observer sees just two at any time. Multiple observers can view the image if enough horizontal viewing freedom is available. For multi-view LCD displays, the lenticular sheet is placed in front of the LCD. For instance, for a five view lenticular display, each pixel in every group of five pixels is directed to a different viewing window (Holliman, 2006). To use the display, five images are sliced vertically into columns and interlaced appropriately. However, multi-view systems need high-speed display elements and high bandwidth image generation and interface circuits that increase their manufacturing cost.

When using MLD technology the user can perceive depth information without needing glasses (e.g. polarised, anaglyph) or mirror setups to separate views. In workplace environments where the use of such additional equipment is either inconvenient or can hinder task execution, this can be an important factor. Because of its architecture, the MLD is cost efficient and could compete with other autostereoscopic parallax displays that are already in the market. Other parallax displays have not been adopted as desktop 3D displays because they are expensive due to the need of high quality optical lenses or the need for high-speed image generation and interface circuits.

2.4. Previous Research on the MLD

The MLD is an innovative technological improvement rather than a solution for an existing problem. As such, very little is known about its usability or the effect of depth in various tasks. So far, the empirical evidence gathered is equivocal.

Wong *et al*. (2005) identified four potential capabilities of the MLD to improve various aspects of visual information search and detection. According to the authors, the MLD supports:

- "Focus + Context" by presenting details on the front screen as the focus of attention and keep the context on the second screen;
- "Visual Linking" by allowing the designer to show relationships between entities in the front and rear screen;
- "Information Layering" by creating physically distinct but visually overlapping information;
- "Information Foraging", based on Pirolli *et al.* theory (1999), by using depth to convey meaning which theoretically increases users' search capabilities (Wong, Joyekurun *et al.*, 2005; Wong, Mansour *et al.*, 2005).

2.4.1. **Focus+Context**

Studies that have analyzed the Focus + Context capability have failed to produce significant results but subjective ratings have favoured the MLD. Bishop (2005) conducted two experiments to evaluate this capability using a 17” MLD. The separation between the screens was not specified. Sixteen students participated in both studies.

A map of the area surrounding the University of Canterbury in Christchurch, New Zealand was presented on the MLD’s rear layer, and a pipe network, coloured in blue, on the front layer. Participants were asked to locate a pipe on a given road and then to locate a break in the pipe which was depicted as a red line perpendicular to the pipe network.

The second experiment presented the same map but it was fixed and semi-transparent. The map was presented on the back layer of the MLD. A pannable zoomed-map on the front layer showed fine details of the roads and the pipe network. A rectangle on the rear layer indicated the current area shown in the front layer. Although only students from the university were recruited and the breaks on the pipes were on streets located near the University to ensure familiarity, geographic knowledge was a confounding factor. Only subjective ratings were published. For the first experiment, participants liked the MLD and they thought it was easy to use and easy to understand. However, for the second experiment, participants were neutral about the MLD and its ease of use, and they rated it as relatively easy to understand.

Hayes *et al*. (2006) also evaluated the Focus+Context capability of the MLD by using a replica of a map display used by an Ambulance Dispatch Centre in New Zealand. The map and secondary information were presented on the rear layer and key pieces of information about ambulance dispatch centres were presented on the front layer. The MLD used for this experiment was a 17” display with 10 mm separation between the screens.

Forty participants, with computer knowledge but no domain-experience, were given an incident location and patient condition. They were asked to assign an ambulance considered to be appropriate for the incident. The incidents were categorized based on four levels of difficulty. Results showed that, overall, there was no significant differences in response times between the MLD and the single-layer condition. An analysis of the four difficulty levels resulted in no significant difference for all levels except level 3 which presented a borderline significance (p=0.053). The authors indicated that fewer errors were made in the MLD but no statistical analysis was reported for accuracy. They also make an interesting observation about one of the trials in which most of the participants made a mistake. Participants chose an ambulance that had the closest linear distance. However, taking into account the road distance, that ambulance was the furthest away from the incident. The authors suggest that this error could have been caused by attentional tunnelling but further research is required to evaluate this observation (Hayes *et al.*, 2006).

2.4.2. **Visual Linking**

Studies of the MLD's Visual Linking capability evaluated whether depth reduces the interference from conflicting stimuli. The "Eriksen" flanker task is one way to evaluate this visual interference. In its original form, an arrow pointing to the left or right is flanked by two distracter arrows creating either compatible (<<<<<) or incompatible (>><>>) trials (Theeuwes *et al.*, 1998). Carr *et al*. (2006) investigated whether the depth separation between non-targets and targets flankers inhibited interference. They recruited 14 participants for this experiment. They used a 15" MLD with 12 mm separation. Results showed that the depth of the MLD did not reduce interference. The authors suggested that the MLD possibly reduces interference when targets and non-targets are in close proximity, but that it may accentuate the distracting effects when larger, full screen eye movements are required.

Another way to evaluate visual interference and the human capacity to direct attention is the Stroop task. The classic Stroop task involves naming the colour ink of words that are either congruent (RED written in red ink) or incongruent (RED written in blue ink). Conflict occurs because people's reading abilities interfere with their attempt to correctly name the word's ink colour.

A Stroop test setup was used by Aboelsaadat and Balakrishnan (2004) to evaluate whether the MLD could reduce interference between overlapping and non-overlapping stimuli. Sixteen participants took part in this experiment. The authors did not specify the display size but mentioned that it has a 14.5mm separation between the screens. The rear layer of the MLD was used for the single layer condition (SLD). A colour name was displayed in the middle of the screen, and in a different colour, a rectangle behind the word (overlapping stimuli) or a line underneath it (non-overlapping stimuli) was rendered. For the double-layer condition, the word was displayed on one layer and, the rectangle or the line, were presented on a different layer. They found that, for overlapping stimuli, performance degraded using the MLD when the stimuli semantically compete for user's attention. For non-overlapping stimuli, the MLD generally equalled the SLD[3] condition.

2.4.3. **Information Layering**

Wong et al. (2005) explored the effectiveness of using depth and alpha-blending to create varying levels of transparency and a sense of visual depth by comparing objects presented on both layers of a MLD to a control condition using a single-layer display. They used a 17" MLD with 14 mm separation. For the easy task, participants were asked to find a circle tagged with a value of 5000. For the hard task, the circles were tagged with three values. Participants had to find a circle with a value one of 5000 and a value two of 10. Ten blue balls were always presented on the screen. In the MLD condition, balls with a value of 5000 or greater were presented on the front

---

[3] Single-Layer Display

layer. The authors did not indicate the lower and upper limits that the values could get. They found that under easy task conditions there was no difference in response times for selecting targets between MLD and SLD conditions. However, in more complex conditions, the MLD showed significantly faster response times than the SLD.

Bishop (2005) conducted another experiment to evaluate whether targets could be selected when stimuli were distributed across two layers. Sixteen students participated in this study. A 10×8 matrix of black aircraft was displayed. Participants had to select 20 targets that were either highlighted by colour (blue); by depth (20 black aircraft rear layer); or by colour and depth (targets: blue aircraft rear layer, distracters: black aircraft front layer). The main screen for the single layer condition was the front layer. Results indicate that there was not a significant difference in response times between the MLD and the SLD. The author concluded that the MLD offers little or no performance benefit over SLD for target selection tasks.

Based on prior findings that information layering in different depth planes improved multiple object tracking (MOT), a study by Bolia *et al*. (2004) tested MOT in the MLD but failed to produce conclusive results. The authors compared depth and transparency in single and dual task conditions (a MOT and a relatively simple digit pair task). In the MLD condition stimuli were displayed on different depth planes. In the SLD condition, transparency was used. The authors did not find an effect of depth on performance. The depth of the MLD showed no benefit over a SLD with some level of transparency. They concluded that transparency alone (in a single layer display) reduces the deficits introduced by occlusion and the perceived mental workload. The authors suggested that the inconclusive results may have been due to the simplicity of the secondary task. They speculated that in a more cluttered simulation or with a more difficult secondary task, depth might be an effective cue.

Dunser and Mancero (2009) also evaluated the effect of depth in a MOT task but combined it with a change detection task. They used a 17" MLD with a 7 mm separation between the LCDs. Twenty participants were asked to track four or six circles from a set of sixteen. The circles were presented on one layer or equally distributed across the two layers of the MLD. For the dual condition, a secondary task was added. One out of four circles located at the four corners of the display changed colour, depth-layer, or colour and depth. Results showed that the MOT condition only was significantly more accurate than the dual conditions. When stimuli were distributed across the two layers, a small but significant improvement in response times was found for the MLD compared to the SLD. For the change detection task, response times were significantly faster for changes in colour compared to changes in depth. The combination of colour and depth was not more advantageous than using colour only. Accuracy for the change detection task was significantly better for changes in colour than changes in depth.

2.4.4. **Information Foraging**

Wong *et al*. (2005) used the Information Foraging concept based on Pirolli's theory (1999). This theory assumes that people, when possible, will modify their strategies or the structure of the environment to maximize their rate of gaining valuable information. Several studies have evaluated the effect of depth in visual search and information retrieval. One of the few applied studies examined performance in information retrieval using a head-mounted display and a MLD (Galster *et al.*, 2006). Eight participants were asked to assess if an aircraft had enough time to attack a target and refuel when using an air battle management system while attending to several secondary tasks. The authors did not indicate if participants had any previous experience with the system. The MLD used for this experiment had 18.1" and 12 mm separation. The control condition was presented on the rear layer of the MLD as a single-layer display. Results indicated that there was no significant difference in accuracy and response times between the different display types. Subjective ratings indicated that participants prefer the MLD.

They rated their workload as lower and their ability to perform multiple tasks as higher when using the MLD.

Another study evaluated the effect of depth in conjunction searches (Dunser *et al.*, 2008). A conjunction search occurs when a target stimulus is defined by a combination of two or more features. For instance, if one searches for an orange square among blue squares and orange triangles, neither the colour "orange" nor the shape "square" is sufficient in isolation to uniquely specify the search target (Palmer *et al.*, 2000) - See section 4.4.1.1 - .

Dunser *et al* (2008) used a 17" MLD with a 7 mm separation between the screens. Twenty participants had to search for a red circle among distracters. Distracters could be blue circles (colour condition); red triangles (shape condition), or red circles on the back layer (depth condition). There were 15, 30 or 45 stimuli. For the control condition, the front layer was used as the SLD. Results show that the fastest search occurred when the red target was presented among blue distracters. The slowest search occurred when the distracters where red circles in the rear layer and red triangles and the target on the front layer. Separating the target only by depth resulted in very slow searches. However, if the target was positioned on the front layer alone, and the distracters (red circles, red triangles and blue circles) were presented on the rear layer, the search slope was flat indicating that the search was done in parallel. While the authors suggested that depth could increase performance in relatively complex searches, their results suggested the opposite. Depth alone was found not to be as effective as colour or shape cues for visual search tasks.

2.4.5. **Design Guidelines for utilising the MLD**

Several promising design techniques have been used to produce applications for the MLD. Unfortunately they have not yet been evaluated but they could be used as a baseline for future research.

Prema et al. (2006) developed rendering techniques for the MLD. They provided some guidelines for producing effective scenes and enhancing perception but did not find a general technique that works well for all applications. These guidelines include emphasizing important objects by displaying them on different layers, separating datasets across different layers, extruding objects across layers, transitioning objects smoothly between layers, and making use of the transparency of the front layer.

Masoodian et al. (2004) developed a word-processing document application for the MLD. Their aim was to enhance the speed at which users could navigate the document by providing constant information about the user's position within the document. The front layer of the display presented a standard editing window with the text shown in full size on a white background. Context was provided by the rear layer which displayed a multiple page preview of the document with a slight reduction in luminance. Although no formal evaluations were conducted, user opinions indicated that the system was easier to use than standard word processing applications.

2.5. Discussion

When working with any three-dimensional display, it is important to know its properties. The MLD produces depth due to the physical separation of the screens but this physical separation cannot be modified.

Like the parallax-barrier and lenticular displays, the quality of the images in the MLD depends on the spatial resolution of the two-dimensional displays used. Although the quality of the images, especially on the rear layer, is one limitation of the MLD model used for this research, it is important to note that it could be improved if the manufacturers used alternatives displays to LCDs.

As is evident from this review, there is little conclusive evidence regarding the MLD's potential impact on change detection tasks. There is only one study that uses change detection as a secondary task (Dunser *et al.*, 2009). The limited empirical investigations have produced equivocal results. While some have reported a positive effect of the MLD usually using visual search or multiple object tracking techniques (Dunser *et al.*, 2008; Dunser *et al.*, 2009; Hayes *et al.*, 2006; Wong, Mansour *et al.*, 2005), others have reported performance degradation (Aboelsaadat *et al.*, 2004; Bolia *et al.*, 2004). Overall, most of these studies included subjective ratings that have generally favoured the MLD (Galster *et al.*, 2006; , 2004; , 2006), even when objective measures have found limited improvement in performance.

Chapter 3

THE HUMAN VISUAL SYSTEM

The MLD is a display that provides a true three dimensional effect to the user by separating images in two distinct layers. Since the characteristics of a display and the nature of human vision combined define the limits of the overall system, it is crucial to have an understanding of the mechanisms that underlie depth perception. This chapter gives an overview of the human visual system especially focused on depth perception.

3.1. The Human Eye

The retina is a light-sensitive layer at the back of the eye that covers about 65 percent of its interior surface (Nave, 2001). Photosensitive cells called rods and cones in the retina convert incident light energy into signals that are carried to the brain by the optic nerve. In the middle of the retina, there is a small dimple called the fovea which is responsible for sharp central vision. The centre of the fovea contains only cone photoreceptors and virtually no rods. There are six to seven million cones concentrated in the centre of the fovea. The high spatial density of cones accounts for the high visual acuity capability at the fovea. Therefore, cones make the largest contribution to the information going to deeper brain centres and provide most of the fine-grained spatial resolvability of the visual system (Duchowski, 2000).

The daylight vision (cone vision) adapts much more rapidly to changing light levels, adjusting to a change of light -like coming indoors out of sunlight- in a few seconds. Like all neurons, the cones fire to produce an electrical impulse on the nerve fibre and then must reset to fire again. The light adaptation is thought to occur by adjusting this reset time (Bruce *et al.*, 2003). The eye moves continually to keep the light from the object of interest falling on the fovea.

The rods, on the other hand, are more numerous than the cones, reaching about 120 million. These photoreceptors are responsible for our dark-adapted, or scotopic vision (Bruce *et al.*, 2003). The rods are incredibly efficient, more than one thousand times as sensitive as the cones. The rod sensitivity is shifted toward shorter wavelengths compared to daylight vision. Because the rod adaption process is much slower than that of the cones, the optimum dark-adapted vision is obtained only after a considerable period of darkness which can take about 30 minutes or longer (Nave, 2001). Since the rods predominate in the peripheral vision, that peripheral vision is more light-sensitive, enabling primates to see dimmer objects in their peripheral vision. Additionally, while the visual acuity is much better with the cones, the rods are better motion sensors, so primates can detect motion better with their peripheral vision (Nave, 2001).

Polyak (1941) subdivided the central retina into three regions, "fovea," "parafovea," and "perifovea." He defined the *fovea* as the area located in the centre of the macula region of the retina. The *parafovea* as the intermediate belt where the ganglion cell layer is composed of more than five rows of cells; and the *perifovea* as the outermost region where the ganglion cell layer contains two to four rows of cells, and where visual acuity is below the optimum. This, in turn, is surrounded by a larger peripheral area that delivers information of low resolution (Iwasaki *et al.*, 1986).

Although Polyak was criticized for his lack of precision in measurements and definitions of what constitutes the inner and outer regions of the retina (Dimmick, 1944), others have based their research about spatial vision and visual acuity on his classification.

The ability of a person or animal to detect fine spatial pattern is expressed as *visual acuity*. It is usually measured by the use of a pattern of parallel vertical dark bars and bright bars with the same width (Bruce *et al.*, 2003). These bars are made narrower to the point that the observer is unable to resolve the grating. Because the perception of the width of the bars depends on how far the observer is from the grating, the width is measured by *visual angle* or the angle that a bar subtends at the eye (Figure 3-1) (Bruce *et al.*, 2003).

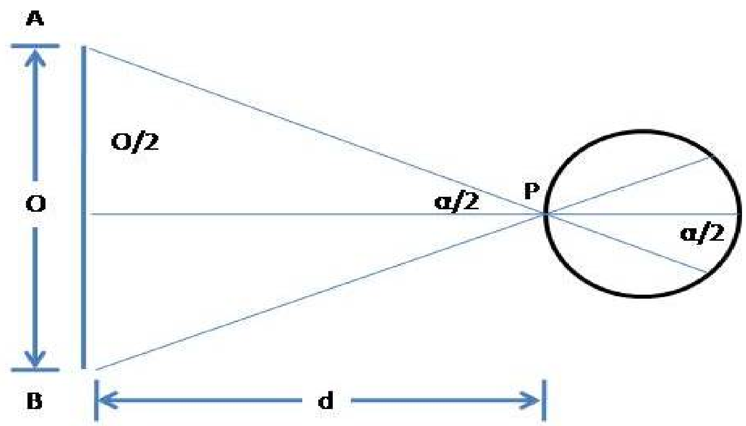


Figure 3-1. Visual angle. Source: (Bruce *et al.*, 2003)

The entire visual field roughly corresponds to a 23400 square degree area defined by an ellipsoid with the horizontal major axis subtending 180° visual angle, and the minor vertical axis subtending 130° (Duchowski, 2000). The diameter of the highest acuity circular region, the fovea, subtends 1.5 – 2° of visual angle, which is roughly equivalent to twice the width of our thumbnail at arm's length (Carrasco *et al.*, 1995). The parafovea comprises 4 to 5° of visual angle and acuity drops off sharply beyond that point. At 5°, acuity is only 50% (Carrasco *et al.*, 1995; Duchowski, 2000).

The useful visual field extends to 30 º of visual angle and anything further than 30º corresponds to the *visual periphery*. Beyond the useful field of view, the rest of the visual field has very poor resolvable power and is mostly used for perception of ambient motion (Duchowski, 2000; Linde, 2003). For the purpose of this research, the region between 5 and 30 degrees of visual angle is referred as *outside the parafoveal* region.

As shown in Figure 3-1, the visual angle, usually denoted θ, is the angle a viewed object subtends at the eye. It is usually stated in degrees of arc. Figure 3-1 shows an observer's eye looking at a frontal extent (the vertical arrow) that has a linear size *O*, located in the distance *d* from point *P*. For present purposes, point *P* can represent the eye's nodal points at about the centre of the lens, and also represent the centre of the eye's entrance pupil that is only a few millimetres in front of the lens. The visual angle θ is the angle between the chief rays for *A* and *B*. The visual angle θ can be measured using the formula (1):

$$\theta = 2 \arctan\left(\frac{O}{2d}\right) \qquad (1)$$

3.2. Eye movements

Sampling of the optic array is achieved by three kinds of eye movement: saccades, pursuit movements and convergence. The rapid and intermittent jumps of eye position called *saccades* are made in order to fixate an object with foveal vision (Bruce *et al.*, 2003). As a person looks at a scene, the eyes make several saccades each second to scan it. Once an object is fixated, *pursuit movements* keep it in foveal vision as it moves, or as the observer moves. However, when the distance of an object from the observer changes, *convergence* movements keep it fixated by the foveas of both eyes. As an object comes closer, convergence movements turn the gaze of both eyes towards the nose, but if the object comes within a few inches of the face, further convergence is impossible and double vision occurs (Bruce *et al.*, 2003).

Saccadic eye movements are very fast, ballistic eye movements separated by fixation periods during which the eyes are relatively still. The term ballistic refers to the presumption that saccade destinations are pre-programmed. That is, once the saccadic movement to the next desired fixation has been calculated (programming latencies of about 200ms have been reported), saccades cannot be stopped or modified (Bridgeman *et al.*, 1994; Duchowski, 2000).

Saccades can reach speeds of up to about 1000 degrees per second. Although every saccade causes a large movement of the image of the environment on our retina, we never perceive this motion. This aspect of perceptual stability is often referred to as saccadic suppression: a reduction of visual sensitivity around the time of saccades (Bremmer *et al.*, 2009).

Visual information is taken in during the fixational pauses between movements. Fixations are defined as eye movements that stabilize the retina over a stationary object of interest (Duchowski, 2000). Therefore, a sequence of fixational pauses provides a sequence of discrete retinal images, a sequence of glimpses or samples separated in time (Bruce *et al.*, 2003).

Additionally, a sharp focused image will only be formed of objects lying at a certain range of distances from the eye because the optical lens and the cornea are both fixed at a specific distance from the retina. This range at which we can see objects is called *depth of field* and is the distance over which the object can move to and from the eye without the image plane falling outside the layer of retinal receptors (Bruce *et al.*, 2003). For a human eye focused at infinity, this range is about six meters to infinity. The reason why we can focus on objects less than six meters from the eye is that its optics can be adjusted by a process called *accommodation* (Bruce *et al.*, 2003).

3.3. Human Depth Perception

A number of mechanisms come into play to enable us to perceive depth. Convergence and accommodation are usually known as oculomotor or 'physiological' cues to depth (Lansdown, 1996). The oculomotor cues are generally regarded as having limited potential to help depth judgement at distances greater than a few meters (Eysenck *et al.*, 2005; Holliman, 2006) but at shorter distances such as a desktop's viewing distance, they become very effective. When working with computer screens, the foveal and parafoveal regions allow fine scrutiny of 3% of the entire screen , assuming a 600 mm viewing distance of a 21-inch monitor (Duchowski, 2000).

In addition to the oculomotor cues, there are several monocular depth cues to create an impression of three-dimensional scenes. These include linear perspective, occlusion and image overlap, texture gradient, shading, and motion parallax. The most powerful monocular cue is motion parallax because it provides the brain with a strong cue to spatial relationships without the use of stereopsis, and is useful whether the motion is produced by the observer, the object in the scene, or both (Holliman, 2006). Nevertheless, the most powerful depth cue is binocular stereopsis. *Binocular stereopsis* is often used for the impression of depth arising from binocular cues (Bruce *et al.*, 2003; Eysenck *et al.*, 2005).

By having forward facing eyes with large overlapping visual fields and a brain that can combine and compare information arriving separately at the two retinas, humans can perceive depth in a way that cannot be achieved without binocular vision (Bruce *et al.*, 2003; Coutant *et al.*, 1993). Binocular vision provides humans with the advantage of depth perception derived from the two slightly different projections of the world onto the retinas of the two eyes. Binocular disparity refers to the difference in the positions and shapes of the images in the two eyes due to the different vantage points from which the eyes view the world (Howard *et al.*, 1995).

Disparity has a magnitude and a sign. A point further than the fixation distance creates an *uncrossed disparity* while points closer to the fixation distance create *crossed disparity* (Figure 3-2). Because disparity decreases with squared distance, the value of stereo vision is greatest in the near space. The disparity in the retinal images of an object decreases by a factor of 100 as its distance increases from 2 to 20 meters (Bruce *et al.*, 2003). Hence, far objects yield disparities that are too small to be detected, and as a consequence, stereopsis is only effective at relatively short distances.

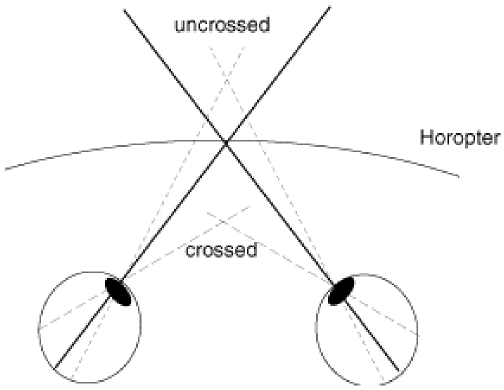


Figure 3-2. Binocular vision – crossed and uncrossed disparity.
Source:http://tinyurl.com/yjz9cue

3.3.1. Integrating cue information

We have seen that depth information is obtained from a variety of sources including oculomotor, monocular and binocular cues. However, if two depth cues provide conflicting information, observers choose one of three strategies: all the information from different depth cues is simply added together; information from a single cue is used and the other one ignored; or the information from different cues interacts in a multiplicative fashion (Eysenck *et al.*, 2005).

Others support, what Helmholtz (1866) called, "unconscious inference" which assumes a probabilistic relationship between cue values and external scene properties. In that case, the task of vision may be seen as to combine these probabilities to derive the most likely set of 3D structures that give rise to the present image (Bruce *et al.*, 2003).

3.3.2. Stereoacuity

Because depth perception in stereoscopic displays involves perceptual integration of multiple sensory cues, it is important to ensure that the observer will find a stereo image pair comfortable to view. Although there are several studies of comfortable perceived depth range on 3D displays, it can be difficult to factor out variables relating to display performance. For reviews on geometric models of perceived depth, see (Holliman, 2004, 2006; Patterson *et al.*, 1994). For tests used to measure stereoacuity, see Appendix I - Stereoacuity Test.

Stereoscopic and some autostereoscopic displays can only generate screen disparity in integer multiples of pixel pitch, therefore the smallest displayable screen disparity is the width of a single pixel. Froner *et al*. (2008) evaluated inter-display depth perception differences to be sure that an observer can detect a screen disparity of one pixel's width, they compared the one-pixel disparity of several 3D displays to measured values of human stereoacuity[4]. However, those measurements do not apply to the MLD, since depth is created by the separation of layers and smaller than one-pixel disparities are possible. Anyhow, the depth difference of the MLD should be compared to values of human stereoacuity.

There are a wide range of values for human stereoacuity depending on the task and the situation in which it is measured. Various studies have shown that the eye is able to see very small values of depth. There have been reports of depth detection thresholds of 1.8 to 2 arc seconds for some individuals (Coutant *et al.*, 1993; Holliman, 2006) which is equivalent to 1.3mm at a viewing distance of 3m. However, this degree of visual precision is achieved by very few people. Coutant *et al.* (1993) indicated that 3% of the population lack or have very poor stereo ability; while Julesz (1971 cf. Coutant *et al.*, 1993) described an informal finding of 2% level of the population completely lacking stereo ability; and Richards (1970 cf. Coutant *et al.*, 1993) found a 4% stereoblindness. In general, we could say that approximately 97% of the population should be able to appreciate depth differences of 2 arc minutes or more in stereoscopic displays (Coutant *et al.*, 1993).

---

[4] Stereo acuity is the ability to detect differences in distance using stereoscopic cues; it is measured by the smallest difference in the images presented to the two eyes that can be detected reliably.

Diner and Fender (1993) suggested that a practical working value for threshold angular disparity can be taken as 20 arc seconds. According to Holliman (2006), a person with a stereoacuity of 20 arc seconds and an eye separation of 65 mm will be able to perceive depth differences between small objects of just 0.84 mm at a viewing distance of 750 mm.

One way to calculate the binocular disparity (δ) associated with a given depth difference (Δ*d*) is given in (2) (Allison *et al.*, 2009),

$$\delta = \frac{\Delta d * e}{Z^2} \qquad (2)$$

This formula indicates that the binocular disparity increases proportionally with the interpupillary distance (*e*) and inversely with the square of viewing distance (*Z*).

It has been difficult to set an exact limit for a minimum viewing distance (*Z*). Optimal eye-screen distances are dependent on the visual capacity, the quality and size of the visual image and the height of the user. Owens *et al.* (1987) found that sustained viewing of visual targets closer to the resting point of vergence contributes to eyestrain. The resting point of vergence is the distance at which the eyes converge when there is nothing to look at, such as in total darkness. It varies among individuals, but averages about 1140 mm when looking straight ahead and 890 mm with a 30° downward gaze angle, although some have suggested that the minimum viewing distance is 630 mm (Ankrum, 1999; Burgess-Limerick *et al.*, 1999). Viewing objects further than the resting point of vergence has not been found to cause any problems (Straker, 2001).

Additionally, mean interpupillary (*e*) distance has been quoted in the stereoscopic literature as being anything from 58 mm to 70 mm. However, it has been shown that it varies with respect to age, gender and race (Dodgson, 2004). Dodgson (2004) suggests that the mean adult interpupillary distance is around 63 mm.

Taking into account that the depth difference (Δd) of the MLD is equal to the 14mm, an interpupillary distance of 63mm and a viewing distance of 630mm, then, using (2), the angular disparity of the MLD due to depth difference is only δ = 8 arc seconds (0.0022°). This binocular disparity is much smaller than the practical working value of 20 arc seconds defined by Diner *et al.* (1993). In order for the MLD to produce a binocular disparity of 20 arc seconds, the viewing distance should be 398 mm which is 1.5 times smaller than the minimum viewing distance of 630 mm.

3.4. Discussion

Psychophysical and physiological studies of depth perception have shown that human vision uses a great variety of information sources, or depth cues, available to the eyes. Individual cues are ambiguous but by combining them, the visual system usually settles on a stable solution. Convergence, accommodation and stereopsis are only effective in facilitation depth perception over relatively short distances, thus, important when dealing with desktop's viewing distances.

The central regions of the retina, the fovea and parafovea, provide most of the fine-grained spatial resolvability of the visual system, while the visual periphery is more sensitive to motion. When working with computer screens, the foveal and the parafoveal regions allow fine scrutiny of 3% of the screen, assuming a 21-inch screen at 600 mm viewing distance.

When dealing with three-dimensional displays, the characteristics of the display and the nature of human vision combined define the limits of the overall system. Depth judgement performance cannot always be predicted from display geometry alone. Other system factors, including software drivers, electronic interfaces, and individual participant differences must also be considered when using a 3D display to make critical depth judgements.

Previous research had established that 20 arc seconds is a practical working limit to use as a value of stereo acuity. The calculations suggest that the

MLD does not produce enough binocular disparity: the angular disparity of the MLD at a viewing distance of 630 mm is only 8 arc seconds. Thus, the MLD produces a very small disparity which means that only a small percentage of the population, those with a high stereoacuity, might be able to easily detect the binocular disparity on the display. Nevertheless, stereopsis is only one of the cues used to evaluate an object's dimensionality. Therefore, when working with the MLD, it might be useful to use occlusion, shading and perspective depth cues since they are very important for any three-dimensional image, and other depth cues such as motion parallax should be combined to produce a stronger depth effect.

Chapter 4

VISUAL ATTENTION AND CHANGE DETECTION

This chapter reviews the issue of attentional control from the perspective of those environmental events that fail to draw attention, particularly *change blindness* and *inattentional blindness*. It assesses previous studies that assume that operators are likely to miss changes if these changes occur while they are multitasking or during an interruption (Di Vita *et al.*, 2004; Durlach, 2004b; Smallman *et al.*, 2003; St. John *et al.*, 2005). This section also reviews previous literature regarding cues that capture attention which may be useful to alert operators to unexpected, infrequent, or high-priority events (Wickens *et al.*, 2008). Finally, this chapter identifies three main gaps in the literature.

4.1. Change Blindness

Change blindness is a cognitive phenomenon that refers to the failure to detect large changes that occur within our visual field. The phenomenon is more likely to occur when the change happens during a visual disruption, (See review Rensink, 2002). A visual disruption masks any transients[5] that would have signalled the change. If this signal is unique or at least larger than the background noise, it will attract attention to its location; if not, the change will not be detected. Without a visual disruption, change blindness occurs if the

[5] A transient is a detectable visual cue that signals a change in the environment over time.

change happens gradually. This gradual transformation suppresses the transient that would have otherwise signal the change (Simons *et al.*, 2000).

Change blindness often comes as a surprise since participants seem to believe that they would have no problem detecting the type of changes presented in laboratory experiments. In fact, when empirically tested, results showed that a large percentage of observers predicted they would detect changes that have not been noticed in previous experimental settings (Levin, 2002; Levin *et al.*, 2002). However, they rated other people less highly. Levin (2002) referred to this metacognitive error as "*change blindness blindness*".

Laboratory experiments in change blindness have used a range of static stimuli ranging from simple shapes like arrays of dots or lines (Phillips, 1974; Rensink, 2000), to more realistic photographs (Curran *et al.*, 2009; Dornhoefer *et al.*, 2002; Pringle *et al.*, 2001; Werner *et al.*, 2000), and dynamic simulations in aviation (Muthard *et al.*, 2002; Muthard *et al.*, 2003; Wickens *et al.*, 2003), naval combat (Di Vita *et al.*, 2004; Smallman *et al.*, 2003; St. John *et al.*, 2005; St. John *et al.*, 2007), military tactical displays (Durlach, 2004a; Durlach *et al.*, 2008) and driving (Most *et al.*, 2007; Pringle *et al.*, 2001) environments. Table 4-1compiles a list of visual disruptions and stimuli used in laboratory experiments and simulations to induce change blindness.

The most common technique to induce change blindness is the flicker technique (Figure 4-1) (Rensink *et al.*, 1997). This technique shows an original image of a scene, a brief blank, and the same image modified in some way (e.g., an item changes colour or is moved). The change is easily seen when the images alternate without intervening blank fields but detection of the change is dramatically impeded if a brief blank of 80 ms or more is interposed between the images. This technique measures the number of alterations required, between the original image and the modified one, for the observer to detect the change. Instead of using the flicker technique, few have opted for a one-shot version (Figure 4-1) in which the modified image is presented only once and observers have to click on the object that they think has changed (Phillips, 1974; Simons *et al.*, 1997b).

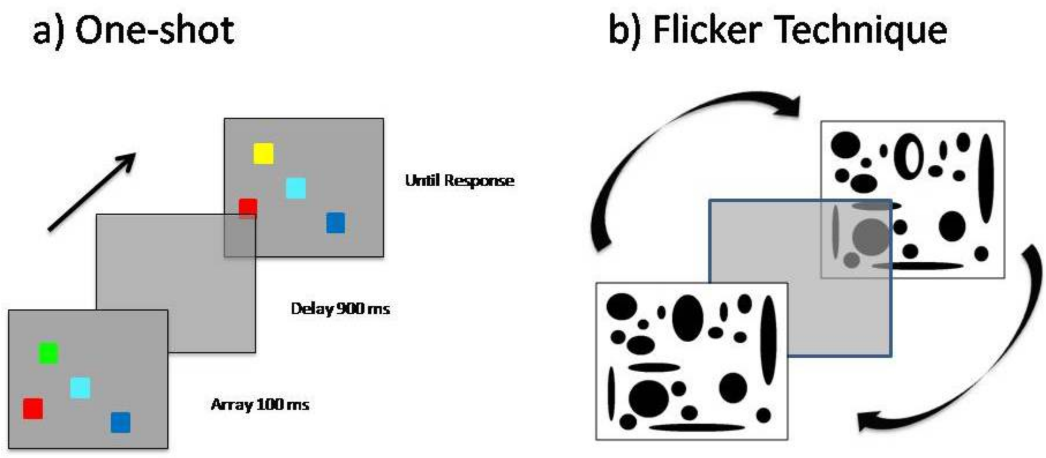


Figure 4-1. a) One-shot technique. b) The flicker technique.

The flicker technique has been adapted to use other visual disruptions besides the blank. For instance, a change could occur during a saccade[6] (Boot *et al.*, 2009; Henderson *et al.*, 1999; Jonides *et al.*, 1982; Kuhn *et al.*, 2010; McConkie *et al.*, 1996; Rayner, 2009), a blink (Dornhoefer *et al.*, 2002; O'Regan *et al.*, 2000; Rensink, 2002), at the same time as the appearance of scattered shapes (Rensink *et al.*, 1999), during an occlusion (Simons *et al.*, 1998), or during a movie-cut (Simons *et al.*, 1997b).

A real-world demonstration showed a high degree of 'blindness' in which an experimenter asked pedestrians for directions. While the pedestrian was providing directions, two additional experimenters, carrying a door, passed between the initial experimenter and the pedestrian. During this brief interruption, a different person replaced the original experimenter. Half the pedestrians continued giving directions after the interruption and failed to notice that they were talking to a different person even though the two experimenters looked different and had distinctly different voices (Simons *et al.*, 1998).

There have been other methods to induce change blindness that do not introduce a visual disruption. Simons *et al* (2000) study presented a series of movies in which the change was presented at 12 frames per second. The change appeared

[6] A saccade refers to the sudden, rapid movements of the eyes. It takes about 100-300ms to initiate a saccade and another 30-120ms to complete the saccade, depending on-among other things-the visual angle traversed (Glenstrup *et al.*, 1995).

to be smooth and continuous. Even without the visual disruption, participants failed to detect changes in the existence of items or colour in photographs of natural scenes. Most experiments conducted with simulations of operational systems have not incorporated visual disruptions (Muthard *et al.*, 2002; Podczerwinski *et al.*, 2002; Wickens *et al.*, 2003) although few have included multiple displays (Di Vita *et al.*, 2004) or general multitasking tasks (Smallman *et al.*, 2003; St. John *et al.*, 2005). The following section reviews studies that have evaluated the change blindness phenomenon using simulations of command and control operational systems.

Table 4-1. Methods used to induce change blindness

Visual Disruption	Stimuli Type	Authors
Blank	Photographs of natural scenes	Rensink *et al.*, 1997; Shore *et al.*, 2000
	Domain-related photographs (i.e traffic related, football related, satellite images)	Curran *et al.*, 2009; Dornhoefer *et al.*, 2002; Pringle *et al.*, 2001; Werner *et al.*, 2000
	Shapes (i.e. rectangles) or drawn-line objects	Austen *et al.*, 2000; Mitroff *et al.*, 2004; Phillips, 1974; Rensink, 2000; Scott-Brown *et al.*, 2000
Saccade	Photographs of natural scenes	Boot *et al.*, 2009; Henderson *et al.*, 1999; Hollingworth, 2003; Jonides *et al.*, 1982; Kuhn *et al.*, 2010; McConkie *et al.*, 1996; Rayner, 2009
Blink	Photographs of natural scenes	Dornhoefer *et al.*, 2002; O'Regan *et al.*, 2000
Scattered shapes	Photographs of natural scenes	Rensink *et al.*, 1999
Occlusion (a door)	Real experimenters and pedestrians	Simons *et al.*, 1998
Pop-up message	Naval monitoring systems and Military tactical displays	Di Vita *et al.*, 2004; Durlach, 2004; Durlach *et al.*, 2008; Smallman *et al.*, 2003; St. John *et al.*, 2005; St. John *et al.*, 2007
Movie-cut	Short videos	Simons *et al.*, 1997
No visual disruption	Motion pictures of natural scenes	Simons *et al.*, 2000
	Aviation simulations	Muthard *et al.*, 2002; Podczerwinski *et al.*, 2002; Wickens *et al.*, 2003; Yeh *et al.*, 2000

4.1.1. **Change Blindness in operational environments**

Some researchers have equated the visual disruption introduced in the flicker technique with interruptions and distractions that happen in the operational environment (Di Vita *et al.*, 2004; Durlach, 2004b). Operators could miss time-critical information on unattended screens if they have to shift their attention between several monitors (Di Vita *et al.*, 2004). Durlach (2004b) affirmed that multitasking can slow change detection. Some have also suggested that in tasks where users monitor dynamic situations, from air traffic management to civil emergency response coordination, interruptions disrupt users' situation awareness and cause them to miss important changes (Durlach, 2004b; Smallman *et al.*, 2003; St. John *et al.*, 2007).

These hypotheses have only been evaluated in laboratory experiments that have examined realistic simulations. When using these part-task simulations, participants' detection rates were poor with detection rates improving when the change was relevant to task and closer to the focus of attention.

Generally, studies in the aviation domain indicate that the relevance of the change to the task being performed is the main predictor of change detection (Muthard *et al.*, 2002; Podczerwinski *et al.*, 2002; Wickens *et al.*, 2003). These studies showed that accuracy improved as the relevance increased. For instance, a study that evaluated several designs of dynamic electronic maps showed that detection of changes that were relevant to the task such as changes that cause a potential conflict while reviewing the flight plan in both traffic and weather systems was faster and more accurate than detection of irrelevant changes (Podczerwinski *et al.*, 2002). In this case, participants were asked to fly north and keep the aircraft at 15000 feet. Overall, changes in heading and airspeed were detected 40% of the time, while only 12% of changes in altitude were noticed (Wickens *et al.*, 2003).

On a similar simulation, Muthard *et al.* (2002) found that changes to aircraft traffic were detected nearly five seconds faster than those to weather systems, suggesting that pilots were visually sampling aircraft hazards more frequently than weather. The authors also noticed that pilots made a plan continuation error by failing to revise their flight plans in approximately one-third of the trials. The authors concluded that change blindness was the cause of these plan continuation errors.

Other factors that also affect change detection are the position and number of objects monitored. Awareness of changes decreased when the number of objects monitored increased (Podczerwinski *et al.*, 2002; Wickens *et al.*, 2003) and changes were located peripherally (Nikolic *et al.*, 2004; Nikolic *et al.*, 2001). Results from these studies highlighted the importance of effective peripheral visual cues and the distribution of tasks and information across sensory channels (Nikolic *et al.*, 2001).

Other researchers have tried to determine the effect of concurrent changes that are typical in a command and control room and the effect of feedback in the detection of changes. Using a command and control system that can be situated in individual vehicles or in tactical operation centres known as the FBC2B, Durlach *et al*. (2008) found that the appearance of military icons were detected faster and more accurately than their disappearance. Detection rates decreased as the number of concurrent changes increased, and providing feedback did not have any significant difference on detection performance. This study, however, recruited students with no prior experience using tactical maps.

Another study that examined change detection in the context of a realistic military command and control station recruited 28 naval Combat Information Centre operators. The study evaluated the operator's ability to detect task-relevant changes in an applied work setting taking into account that critical events in tactical situation displays are often temporal, and spatial antecedents render the change logical and tactically meaningful (Di Vita *et al.*, 2004). Participants were asked to monitor map activity of 8 aircraft and 8 vessels. There were 20 critical changes equally distributed into four categories of attributes changes: course, speed, range or bearing. Participants were notified that a change had occurred and instructed to click on the object that changed in any of the four attributes. They were given feedback after each click. The accuracy of change detection was plotted against the number of selections. If the operator had not chosen correctly by the third choice, statistical modelling indicated that subsequent selection was no better than chance. Results showed that almost 30% of the critical changes required at least two selections to be correctly identified and approximately 15% of the critical changes were selected by pure chance. The authors believe that their results represent a conservative estimate of the magnitude of change blindness in naval environments. Participants were prompted after the change and the scenarios were simpler than real displays which usually present 50 to 100 objects of interest (vessels and aircraft).

Others have attempted to develop change detection tools to support operators recover situation awareness after an interruption. One of these tools is called CHEX (Change History EXplicit) that uses a naval system intended to monitor airspace activity. CHEX automatically detects and logs changes into an interactive table (Smallman *et al.*, 2003). It provides a table next to the tactical map where changes are logged. Highlighting one of the table entries highlights the respective aircraft or vessel icon in the map. The last version of CHEX was evaluated against an Instant Replay tool. The Instant Replay proved worse than no support; but detection of relevant changes to air traffic was faster and more accurate when CHEX was available (Smallman *et al.*, 2003; St. John *et al.*, 2005; St. John *et al.*, 2007). Participants for these studies were recruited locally and were not subject-matter experts.

The results of these studies have shown that operators can miss important changes even when they are not fatigued, stressed or multitasking. They suggested that the design of future military, aviation, or naval digital display systems must take into account human limitations in detecting visual changes especially if multiple events occur concurrently, or operators are often interrupted.

Although some of these studies have not recruited subject-matter experts, they have been successful in replicating the phenomenon of change blindness using realistic simulations. Some have suggested that the magnitude of the problem could be worse in the operational context since these simulations have simplified the scenarios or participants were prompted about the occurrence and type of changes.

These studies have determined some factors that reduce but do not eliminate change blindness. With simpler stimuli, results have shown that detection accuracy declines as the number of objects increased (Phillips, 1974), and that deletion of an item appears to be detected more easily than its addition (Rensink, 2002). With more complex scenes, significance of the change, relevance of the change to the task, domain expertise and familiarity with the scene usually provoke higher detection rates. Results have demonstrated that change detection depends greatly on the significance of the part of the scene being changed, with faster and more accurate identification for those structures of greatest interest (O'Regan *et al.*, 2000; Rensink *et al.*, 1997). Detection rates increase if the change is relevant to the task, for instance, a pilot will notice an aircraft if it conflicts with his/her flight path (Podczerwinski *et al.*, 2002). Expertise on the domain has also shown to have a positive impact on detection but experts are still prone to missing changes (Curran *et al.*, 2009; Werner *et al.*, 2000). Finally, familiarity with the objects and scenes seem to increase the rates of detection (Archambault *et al.*, 1999; but see Rosielle *et al.*, 2008).

4.2. Inattentional Blindness

Inattentional Blindness has been described as the *looked-but-failed-to-see* effect (Herslund *et al.*, 2003). It refers to the failure of seeing an unexpected event within the visual field when the observer’s attention is diverted to a primary task (Mack *et al.*, 1999).

Most of the evidence for Inattentional Blindness comes from relatively simple laboratory tasks. However, recent evidence suggests that talking on a mobile phone while driving dramatically increases the probability of hitting an incoming vehicle such as an unexpected bicycle (Herslund *et al.*, 2003). These studies suggest that the more people focus on aspects of their visual world other than the detection of unexpected objects, the less likely they are to detect such objects (Mack, 2007; Scholl *et al.*, 2003).

Studies of change blindness assume that, with attention, features can be encoded and retained in memory (Rensink, 2002), suggesting that all of the information in the visual environment is potentially available for attentive processing. Yet, without attention, not much of this information is retained across views (O'Regan *et al.*, 2000; Rensink, 2002). However, studies of inattentional blindness have claimed that, without attention, "observers may fail not just at change detection, but at perception as well" (Simons *et al.*, 1999).

Mack and Rock (1999) coined the Inattentional Blindness term when using a technique that guarantees that the observer would neither be expecting nor looking for the object of interest. Observers were asked to report the longer arm of a cross which was located at fixation or in the parafovea within 2.3° of fixation. On the third and fourth trial, a "critical stimulus" was presented without warning along with the cross. The third trial was called the "critical inattention trial", and immediately after this trial, participants were asked whether they had seen anything other than the cross. The critical trial was extremely important to obtain data regarding perception without attention because participants were not searching for the specific stimulus and did not anticipate it. The subsequent trials were explicitly divided attention trials because participants were asked to report the longer arm of the cross and anything else that might be present in the screen. A final set of trials, or the full attention control trials were also introduced at the end of the session in which participants were asked to ignore the cross and just report if anything else was presented on the screen.

A main disadvantage of this method is that it permits only one true, critical inattention trial per subject, because for subsequent trials, subjects are likely to be expecting the critical stimulus to appear. Therefore, a large number of participants are required.

Prior to being named *Inattentional Blindness*, a dynamic version of this phenomenon was demonstrated by Neisser and Becklen (1975). Observers viewed a display which presented two overlapping, partially transparent simultaneous videos. The first one presented two people playing a hand-slapping game while the other video presented three people passing a basketball. Participants were asked to monitor one of the two events. In the 7^{th} and 8^{th} trial respectively, the two hand-game players shook hands, and one of the basketball players threw the ball out while the players pretended to continue the game. Only 50% of participants detected the unexpected events but presented difficulty reporting these events accurately (Neisser *et al.*, 1975; Simons *et al.*, 1999). In subsequent studies, Neisser used different versions of the basketball game task. In one of these versions, he superimposed a video showing a team of people wearing black shirts and a second video showing another team wearing white shirts. Participants were asked to monitor one of the two teams and count the number of passes. After 30 seconds, a woman with an open umbrella walked across the screen being visible for 4 seconds. The game continued for another 25 seconds after the appearance of the woman. Only 21% of the participants reported seeing the woman (Neisser and Dube 1978 cf. Most *et al.*, 2001; Simons *et al.*, 1999).

Later on, Simons and Chabris replicated this experiment but without superimposing videos. One video was shown which lasted 75 seconds. It presented two teams of people, one wearing black and the other wearing white shirts, passing a basketball. A woman in a gorilla costume or a woman with an open umbrella walked through, staying in sight for 5 seconds. Overall, only 42% of the observers (n=192) noticed the unexpected event (Simons *et al.*, 1999). However, 58% of the participants monitoring the black team detected the gorilla, but only 27% of those monitoring the white team detected it. Their results suggested that it is more likely to notice an unexpected object that is similar to other objects in the display. The likelihood of noticing an unexpected event was also affected by the difficulty of the monitoring task.

4.3. What have Change and Inattentional Blindness studies achieved?

Both change blindness and inattentional blindness have shown that attention plays a critical role in perception and in representation. These two similar phenomena have demonstrated that without attention, we often do not see unanticipated events, and even with attention, we cannot encode and retain all the details of what we see.

Research on change and inattentional blindness has helped in articulating concepts that were not well understood and has inspired claims on theories of visual attention, visual memory and visual awareness.

The studies mentioned in the literature have helped clarifying the distinction between motion, change, and difference perception. Motion perception refers to the detection of unorganized flow at a location (Simons & Rensink, 2005). Change perception refers to the detection of an ongoing transformation of a structured object, becoming a variation referenced to a structure (Rensink, 2002). Difference perception refers to an inferential comparison of the current stimulus with traces in long-term memories (Rensink, 2002; Scott-Brown *et al.*, 2000; Simons & Rensink, 2005).

Change and Inattentional Blindness findings have also challenged traditional theories of perception, such as the theory that our memory accumulates the contents of successive eye fixations building a complete internal model of a scene (see Bridgeman *et al.*, 1994; Deubel *et al.*, 2002; Irwin, 1991). Based on this evidence, some have argued that change blindness implies that internal visual representations are completely absent (O'Regan *et al.*, 2001); others have argued that successful change detection requires both a representation of the scene before the change and after the change (Scott-Brown *et al.*, 2000; Simons & Ambinder, 2005); while others have suggested that change blindness implies that our representations of visual scenes are sparse or incomplete (Rensink *et al.*, 1997; Simons *et al.*, 1997a).

Probably change blindness research's most important finding is that change detection is mediated by attention. Without focused attention, we cannot perceive changes (Rensink, 2002). In this view, attention must be directed to the region of space in which a change occurs at the time the change takes place. Despite the evidence that suggests that attention is required to detect change, some have argued that although attention may be focused on an object, blindness to large changes may still occur (Austen *et al.*, 2000; Simons *et al.*, 1997b; Simons *et al.*, 1998). This failure could imply limitations on the comparison mechanism used for change detection (Scott-Brown *et al.*, 2000; Simons & Ambinder, 2005), or limitations on the capacity of attention (Simons & Ambinder, 2005) with more evidence supporting the latter.

Even with a limited capacity, research has shown that attention can be guided and controlled. Studying what captures attention could probably help designers to use effective cues to guide operators' attention to changes that could potentially be missed if attention is not drawn to the change information.

4.4. If attention is needed to detect change, can change blindness be mitigated by attentional manipulation?

Attention plays an important role for the detection of expected changes and unexpected events. Previous research has shown that attention can be captured. It seems that researchers have concurred with an interaction between goal-driven attentional control and stimulus-driven attentional capture.

4.4.1. **Features that capture attention**

This section reviews the literature on visual alarms, alerts and cues that capture attention and the effect of their location on attention capture.

4.4.1.1. Visual cues

According to Wickens (2008), *visual search* is one of our most common and important attentional skills because it not only pervades everyday behaviour but it is also a critical component of many specialized tasks such as driving, map reading, medical image interpretation, baggage x-ray screening, menu search, to name a few.

Visual search is probably the most widely technique used for assessing how goal-driven and stimulus-driven selection occurs (Wickens *et al.*, 2008). By definition, visual search tasks require observers to scan the visual environment for a particular object or feature (the target) among several distracter elements. The location of the target is not known a priori. The observer has to indicate whether a search target is present or absent. Reaction times are measured. The measure of attention in the search task is often manifested as a slope of the response time function over the number of distracters (Quinlan, 2003; Treisman *et al.*, 1980; Yantis, 1998). If this slope is flat, it means that the visual cue is processed in parallel and therefore preattentively. A flat slope means that regardless of the number of items in a display, the target detection time stays the same. Based on this result, the Feature Integration Theory states that the processes that underlie human vision are often divided into two fundamentally different classes: operations that are carried out in parallel and operations that are processed serially, dividing vision into an early *preattentive* and a subsequent *attentive* stage (Treisman *et al.*, 1980).

For the design of visual displays, this division is extremely important because visual cues that are process preattentively are almost automatically detected therefore yielding a faster and more natural way of acquiring information. As Kosara (2002) stated one very important aspect of any visualization is that “it utilizes one of the channels to our brain that have the highest bandwidths: our eyes. But even this channel can be used more or less efficiently” (Kosara *et al.*, 2002). One way to use it efficiently is to use *preattentive processing* (Kosara *et al.*, 2002). *Preattentive processing* of visual information, according to Treisman (1985) is performed automatically on the entire visual field detecting basic features of objects in the display and Healey (2005) suggests that tasks that can be performed on large multi-element displays in less than 200 to 250 milliseconds are considered preattentive.

According to the Feature Integration Theory, colours, closure, line ends, contrast, tilt, curvature and size are simple features that are extracted from the visual display in the preattentive system and later joined in the focused attention system into coherent objects. She argues that preattentive processing is done quickly, effortlessly and in parallel without any attention being focused on the display (Treisman, 1982, 1985; Treisman *et al.*, 1980).

According to Wolfe, there are large amounts of convincing data demonstrating that colour, motion, orientation and size are features that are processed in parallel and therefore preattentively (Wolfe, 2007; Wolfe *et al.*, 2004). He indicates that evidence for other attributes such as curvature, novelty, luminance onsets, vernier offsets, shape, and transparency could guide attention under certain conditions but the evidence is not conclusive (Wolfe *et al.*, 2004). For instance, in the case of stereoscopic depth, Wolfe and Horowitz (2004) suggested that there might be a broader dimension of something like three-dimensional layout that would capture various depth cues including stereopsis, the various pictorial depth cues, and shading. The cues would merely serve to create three-dimensional surfaces in the way that wavelength (not a guiding dimension) creates colour.

The Feature Integration Theory also describes feature and conjunction searches. The former is the process of searching for a target which differs from the distracters by a unique visual feature, such as colour, size, orientation or shape and is processed in parallel. For instance, a red target is found quickly if it is among blue distracters. A conjunction search, on the other hand, is the process of searching for a target that is defined by the combination of two or more features and therefore is processed serially. For example, when locating an orange square among blue squares and orange triangles, one must integrate the colour (orange or blue) and the shape (square or triangle) feature for every separate item in the display until the target is found (Treisman, 1982; Treisman *et al.*, 1980).

Nonetheless, in the late 1980s, Nakayama and Silverman argued that some conjunction searches could be performed in parallel (Nakayama *et al.*, 1986). They found that conjunction tasks combining stereoscopic disparity of 20 arc min with either colour or motion were qualitatively different and much easier than other conjunctive searches. They argued that the visual system can perform a parallel search in one depth plane without interference from target-like distracters in another depth plane (Nakayama *et al.*, 1986). O'Toole and Walker (1997) determined that stereoscopic depth (4 arc min) provoked efficient searches only if it was combined with another feature, but if used alone, stereoscopic depth produced inefficient searches. In fact, Theeuwes *et al.* (1998) found that when targets and distracters were identical in colour and just separated by stereoscopic depth (25arc min), objects in different depth planes slowed the search. However, when the colours of the target and distracters were different, response times were a lot faster concluding that directing attention to a particular depth plane can prevent attentional capture from another depth plane (Theeuwes *et al.*, 1998).

Dunser *et al.* (2008) conducted a visual search study using conjunction searches that manipulated colour, shape and depth. They used a 17" MLD with 7 mm separation between the screens but did not specify the binocular disparity. Their results showed that not only participants were able to deploy attention in different depth planes, but also were able to detect the target if it was separated in depth from the distracters. The slope of the response time function for the depth cue was flat but at a longer latency than colour, similar to the results obtained by Nakayama (1986). Because of the longer latency, it could be argued that depth was not processed in parallel. Their results also suggested that if depth is combined with colour and shape, the detection of the target was almost as fast as colour.

Nevertheless, De la Rosa (2008) found that the efficiency of the search depends on the stereoscopic disparity. While the stereoscopic display used by Nakayama and Silverman presented a 20 arc min binocular disparity, O'Toole's one presented a binocular disparity of only 4 arc min. The authors concluded that observers need a separation of about 6 to 7 arc min for the search to be effective and minimise the intrusion of distracters from one plane onto another.

4.4.1.2. Exogenous and endogenous cues

Many have compared the effects of endogenous and exogenous cues. Endogenous cues denote the target location symbolically but not appear at the location, for example an arrow, or the words *left* or *right (Pylyshyn, 2006)*. Exogenous cues are transient signals that appear at the location to which attention is to be shifted, for example a luminance change (Wickens *et al.*, 2008).

Researchers have found that characteristics of the cue stimulus itself influence attentional performance. The first to demonstrate this was Jonides who used a visual search task combined with a cue-validity procedure in which participants viewed an array of eight letters arranged in a circle so that each letter in the array was equidistant from the fixation point (Jonides 1981 cf. Pylyshyn, 2006; Wickens *et al.*, 2008). Participants were asked to press a right button if an R was presented or a left button when an L appeared in the display. An arrowhead would appear before the array indicating one of the locations. The arrowhead could appear either at fixation (endogenous cue) or near the letter location that it indicated (exogenous cue). Jonides demonstrated that exogenous cues tend to draw attention even if they do not predict the target location with accuracy better than chance.

Yantis and Jonides (1984) used a visual search task to determine whether an abrupt onset captures attention automatically and manipulated the number of items in the display. They found that regardless of the number of items in the display, the response times when the target was an abrupt onset did not increase, but, without the onset, the response times to find the target increased significantly with display size. The flat slope obtained from the response times of onset targets strongly suggests that the onset stimuli was identified first during the search and therefore captured attention.

Some have argued that abrupt onsets capture attention only if the observers are set to look for them (Mulckhuyse *et al.*, 2008). Others have found evidence that even when observers have an attentional set for a colour singleton, an irrelevant new object presented with an abrupt onset interfered with the search. These results suggest that abrupt onsets or new objects appear to capture attention independently of top-down control settings (Schreij *et al.*, 2008; Theeuwes, 1994, 1995).

Other work in this area has found that attention responds more quickly to exogenous cues but present a transient response fading of 100–300 ms after cue onset while the more persistent endogenous cues (i.e. arrows) came into effect after about 350 milliseconds (Chastain *et al.*, 1999; Engbert *et al.*, 2003; , 1989; Theeuwes, 1994; Yantis *et al.*, 1984, 1990).

4.4.2. **Attention guidance for interface design**

4.4.2.1. Visual alarms, alerts and abrupt onsets

The nature of the funding focused the research to evaluating the depth of the MLD as a visual cue to highlight changes. Although auditory stimuli are the most reliable attention grabber for alarm systems (Wickens *et al.*, 2008), this section will not review them because the focus is on visual cues. Additionally, in operational environments, operators are already loaded with voice communications. Therefore, their auditory channel is already saturated. This section will review visual alarms that could be used to guide operators' attention to visual changes.

Colour and flashing are two important and common means of coding visual alerting signals. Colour is particularly useful for memory coding and message recognition (Chan *et al.*, 2009). Flashing, on the other hand, if used as a redundant cue, has been found to be superior to colour alone in attracting attention to objects in a display (Chan *et al.*, 1997; Thackray *et al.*, 1991) and in influencing delectability of signal from a distance (Chan *et al.*, 2007).

A direct generalization from basic research on attention capture to alarm and alert design is that onsets tend to capture attention (Wickens *et al.*, 2008). Hence, the most effective visual alarm will be the flashing signal because it entails repeated onsets, any of which may eventually be noticed as the eyes are busy scanning the environment (Wickens *et al.*, 2008). Previous research examining the detection of multiple transient changes within a simulated sonar display demonstrated that performance suffered when observers were asked to detect more than two or three targets but the presence of a flashing cue greatly attenuated this performance deficit even for short flashing-durations (one onset) (Boot *et al.*, 2007).

A second generalization is that unique colours can be effective as alarms and alerts but the ability to capture a singleton will be reduced if nearby stimuli are also colour coded. It is easy to find a singleton if it is more salient than the distracters. For instance, if the participant is asked to detect a red target, it will "pop out" if the red target is among blue distracters. However, the salience of a uniquely coloured item decreases if the background stimuli presents heterogeneous colour (Wickens *et al.*, 2008).

4.4.2.2. Visual cues as filtering tools

Colour, luminance, and flashing have been used as filtering tools to direct the observer's attention to a subset of items assumed to be the most relevant for the task.

Previous research in computer displays has evaluated several types of highlighting attributes for menu search (Fisher, Coury *et al.*, 1989; Fisher & Tan, 1989). They found a significant advantage of highlighted displays over non-highlighted ones as long as the validity of the highlighting cue is greater than 50%. The authors concluded that colour had a significant advantage over flashing and reverse video. The authors assumed that flashing took considerably longer because the flashing target could not be identified during the off portion of the flashing cycle.

Research on the design of other displays has shown that clutter slows visual search, for instance, the time needed for a controller to detect an air-traffic conflict increases proportionately with the number of aircraft in the display (Remington *et al.*, 2000). A way to enable efficient search despite the high levels of clutter is to encourage attentional filtering based on visual features (Wickens *et al.*, 2008).

Yeh and Wickens (2000) evaluated colour and intensity coding in an electronic map display. They compared the results with a third technique that they called "flexible decluttering" in which participants could remove and recall an entire set of information manually. The maps presented terrain features, river and roads, and travelling troops and stationary units. For the colour-coded condition, all three levels of information were depicted in unique colours. The intensity coding highlighted either the river and roads, or the troops and units and lowlighted the rest of the information. For the flexible decluttering, the information was presented in the same colour and intensity, but while the terrain feature was always visible, participants could remove the other two levels of information.

The results suggest that highlighting was highly successful allowing the presentation of information without hindering the focus of attention on one information domain or another. The ability to remove layers of information presented a significant cost when the information needed to be retrieved increasing response times between 2 and 4 seconds. Their results also suggest that colour coding allowed one second faster acquisition of target objects. They concluded that although intensity coding may be less effective than colour, it nonetheless can aid visual search (Kroft *et al.*, 2003; Wickens *et al.*, 2004; Wickens *et al.*, 2003; Yeh *et al.*, 2000).

4.4.2.3. Visual cues that allow comparison

The use of visual cues to present information in a way that is easier to compare and draw relations between the objects has mostly been studied with air-traffic control radar displays. Previous work with air traffic control radar-like displays has shown that the addition of perceptual cues can improve relational comparisons. Results have indicated that coding aircraft altitudes via colour differences improved conflict detection, and the ability to identify aircraft that were within 1000 feet of a target aircraft (Palmer *et al.*, 2008; Remington *et al.*, 2000).

Other efforts have focused on improving relational comparisons by using depth to represent altitude. Pictorial depth cues and stereoscopic depth have shown that air traffic controllers form a more robust 3D mental representation of air traffic. Three-dimensional images compressed on a flat screen leave all three dimensions relatively uncertain due to the distortion of distances and angles, which make three-dimensional displays poor for precise relative position tasks (Ellis, 1985; St. John *et al.*, 1999; St. John *et al.*, 2000; Wickens *et al.*, 1993).

Recent research has explored the potential benefits of adding the visual cues of size and contrast to air traffic displays to aid the apprehension of aircraft altitude, and to detect possible conflicts (Palmer *et al.*, 2008). The authors argued that relative size is an effective cue to indicate depth when stereoscopic and motion cues are absent, and likewise, variations in the relative contrast of aircraft icons may also lead to the impression of depth segregation. For their experiments, aircraft with higher altitudes were portrayed as being larger; and those aircraft at lower altitudes were dimmed assuming that they are farther away from the observer. The authors suggest that size and contrast are features to which the visual system can efficiently guide attention for making relational comparisons. The authors found beneficial effects when both cues, size and contrast, were combined showing an increase in accuracy and response times for the detection of traffic conflicts.

Other research has explored the relation between visual search and change detection using comparison techniques. Comparison techniques (based on the spot-the difference game) present the stimuli either side by side or in sequence. Stimuli usually comprise pairs of images, of which, some differ in colour, size, shape or category (i.e. a cat belongs to the animal category and a sofa belongs to household furniture). Participants have to judge if the two stimuli are identical or different. Overall, results have shown that determining that the images are different is faster than judging them identical (Belke *et al.*, 2002; Brunel *et al.*, 1997; Farell, 1985; Hyun *et al.*, 2009; Meiran *et al.*, 2002; Ninio, 1998, 2004).

Hyun *et. al* (2009) found that the number of objects in a display influence response times more strongly in the absence of a change than in the presence of a change in the same way that the response times slopes in visual search experiments are steeper when the target is defined by the absence of the feature than when the target is defined by the presence of the feature.

Another study that combined comparison and change detection techniques manipulated spatial configuration of a visual display to determine whether people encode spatial arrangements. Results showed that as long as the overall configuration of the display is preserved, participants are more likely to report the change regardless if the configuration is irrelevant to the task (Boduroglu *et al.*, 2009).

4.5. Gaps in the literature

The evidence from change and inattentional blindness has demonstrated that we are blind to changes if we do not attend to them, and even when attending to specific locations, sometimes we fail to see unexpected events. Based on these findings, some have argued that change blindness is a problem in operational environments, but this hypothesis has only been tested in the laboratory.

The first gap in the literature is the lack of research in real context-specific environments with domain experts, and not only with part-task simulations in the laboratory. Some of the studies mentioned above were successful in replicating the change blindness phenomenon using realistic simulations, but their conclusions have been generalized even when the participants did not have prior experience with tactical maps.

Second, results from studies that have evaluated whether depth is a preattentive cue are equivocal. Regarding the use of visual cues to attract attention, there is convincing evidence that colour, size, motion, orientation and abrupt onsets are very powerful cues. Nevertheless, there are some concerns about the use of colour and flashing in the design of interfaces for operational environments where both cues are probably overused. Generally, research on visual onsets and singletons has been conducted using a single monitor, but operators in command and control rooms usually work with multiple monitors. The results from the research mentioned above cannot be generalized to these environments and therefore, one must acknowledge that abrupt onsets and singletons do not invariably capture attention. Additionally, visual alerts could block information in the background. For instance, a soldier wearing a head-mounted display who receives a visual alert to a possible enemy target location, might have difficulty discerning whether the target is actually an enemy because the visual alert dims or blocks the location (Wickens *et al.*, 2009; Wickens *et al.*, 2008).

The use of depth to capture attention has been evaluated but results from these studies are not conclusive. Nevertheless, recent research has demonstrated that the efficiency of the search when using depth as a cue depends on stereoscopic disparity. Although stereoscopic depth alone seems to be processed serially, when combined with another feature, such as colour or motion, and a binocular disparity greater than 6 arc min, conjunction searches are processed in parallel. Further research is needed to determine whether depth can be used as an alerting tool.

Finally, comparison techniques have been used to explore the relation between visual search and change detection. They have also been used to explore whether people can detect relations within the information presented if this information is colour, size or intensity coded. Some of these visual cues have been used to depict monocular depth cues. However, binocular depth has not been used for these specific tasks and, hence, the need for further research to examine whether depth is an effective way to allocate attention in different depth planes to enhance the detection of differences.

Chapter 5

EVALUATING THE MLD AS AN ALERTING CUE

This chapter describes two laboratory experiments set up to evaluate whether depth has an effect on detection when used as a cue to guide attention to expected and unexpected changes.

5.1. Introduction

The literature suggests that effective display design requires cues to guide attention to crucial information (Wickens *et al.*, 2008). Previous studies on visual search had shown compelling evidence about the efficacy of colour, motion, orientation and size to direct attention (See Chapter 4 and reviews Fisher, Coury *et al.*, 1989; Treisman *et al.*, 1980; Wolfe, 2007). Thus, current practice in conventional display design uses colour or flashes as visual cues to highlight and attract attention to the changed data value.

The first experiment evaluated detection when the change is expected and attention is focused on the task with no visual disruption, similar to Simons *et al* study (2000). The second experiment evaluated detection of an unexpected event which meant that participants' attention had to be diverted to a secondary task while maintaining the detection task unknown. This second experiment replicated Most *et. al* (2000) study on Inattentional Blindness including depth as a variable.

Ethics Approval for all experiments was obtained from Middlesex University Ethics Committee and the AFRL Institutional Review Board (See Appendix II – Ethics Approval).

5.2. Experiment 1: Is the MLD's depth a sufficient alerting cue?

This experiment evaluated whether depth is a sufficient visual cue to highlight expected changes. The participants' task was to detect a change, and therefore, the change was expected. No visual disruption was included like in Simons *et al* study (2000) but the duration of the visual transients was manipulated. A transient is a detectable visual cue that signals a change in the environment over time. Because there is no visual disruption, this experiment evaluates what Rensink (2002) defined as dynamic changes (See section 1.3).

A depth and a colour transient were used to highlight changes. These transients were compared to a non-transient condition. Transient durations were manipulated to 250, 350 and 450 ms. These rates were based on previous studies on cues used to orient attention in visual space that found that exogenous cues, such as sudden flashes, had a powerful transient response that fades 100–300 ms after cue onset, while endogenous cues, such as arrows came into effect between 350–600 ms (Engbert *et al.*, 2003; , 1989). Eccentricity of the changing stimuli was also analysed to assess the effect of depth and spatial location in the detection of changes (see section 4.4.1.2).

It is important to note that the main objective of Experiment 1 was to evaluate whether depth was a sufficient visual cue for change detection. If depth happened to be better than the non-transient condition, then the effect of depth could be compared to colour. Colour was chosen because it is the most widely visual cue and there is enough evidence to suggest that colour is a preattentive cue (See section 4.4).

5.2.1. **Materials and Methodology**

5.2.1.1. Pilot study

A pilot study was conducted prior the main study. Five student researchers volunteered. The pilot study examined whether the instructions given were clear, the experimental design was complete, the data collected was correct and the time allocated per block was suitable. The average task execution time was calculated prior the pilot study and tested during it.

To calculate an appropriate time limit, the task execution time was estimated using the Keystroke-Level model (Card *et al.*, 1983). The KLM requires only that the user interface be specified in enough detail to dictate the sequence of actions required to perform the tasks of interest (Kieras, 2001). According to Kieras (2001), the average operator requires 1.1 seconds to point the cursor to a desired place on the screen, 0.1 seconds to click or release the mouse button, and 1.2 seconds to perform a routine cognitive process which in this case is detecting the change. Although the experimental procedure will be explained in more detail in section 5.3.1.4, for now, it suffices to say that participants required to look at the fixation cross while being attentive to any possible change (M), point the mouse to the digit that changed (P), press the button (B), and finally come back to the fixation cross (P). Thus, the estimated total time was calculated by adding each of the actions: M + 2P + B = 1.2 + 2(1.1) + 0.1. This lead to a time limit of 3.5 seconds and a delay between 1000 and 2000 milliseconds was added to account for slower users.

With the inclusion of a time limit, and a few minor corrections to the underlying codebase and the instructions, a preliminary analysis of the data was performed. Results from the preliminary analysis suggested that changes located outside the parafoveal region were detected more accurately when highlighted with a depth-transient. Results from studies reviewed in Chapter 4 suggested that colour, motion, orientation and size are effective cues to guide attention (Fisher, Coury *et al.*, 1989; Treisman *et al.*, 1980; Wolfe, 2007) whereas stereoscopic depth and pictorial depth cues have shown ambiguous results (Holliday *et al.*, 1991; Nakayama *et al.*, 1986; Theeuwes *et al.*, 1998). Based on these points, three hypotheses were tested:

1. Detection of changes highlighted with a colour-transient would be more accurate than those highlighted with depth.
2. Detection of changes highlighted with a colour-transient would be faster than those highlighted with a depth-transient.
3. Detection of changes highlighted with a depth-transient would be more accurate when presented outside the parafoveal region.

5.2.1.2. Participants

A total of 22 participants, 21 Middlesex University students and one staff member, took part in this experiment. Thirteen were males and nine females with a mean age of 23 years (SD=5) varying from 18 to 40. Participants received a voucher for a local store. All participants reported normal or corrected-to-normal vision. Each observer passed the Ishihara test for colour blindness. The test was conducted online before starting the experiment (http://tinyurl.com/26wm3c). Stereoacuity was not measured.

5.2.1.3. Equipment

Stimuli were presented on a 17-inch Multi-Layered Display (MLD) with 14 mm separation between the screens. The MLD was set at a resolution of 1024×768 per screen.

5.2.1.4. Stimuli

A 10×8 matrix of digits was presented on the back layer of the MLD. The matrix presented only single digit numbers. The digits were blue against a white background. The digits were separated horizontally and vertically by 90 pixels. The font used was Arial 20 pts. Figure 5-1 shows a screenshot of the matrix. To analyze eccentricity, the screen was divided in visual regions. The labels and dotted lines were not presented during the experiment but assisted in the analysis.

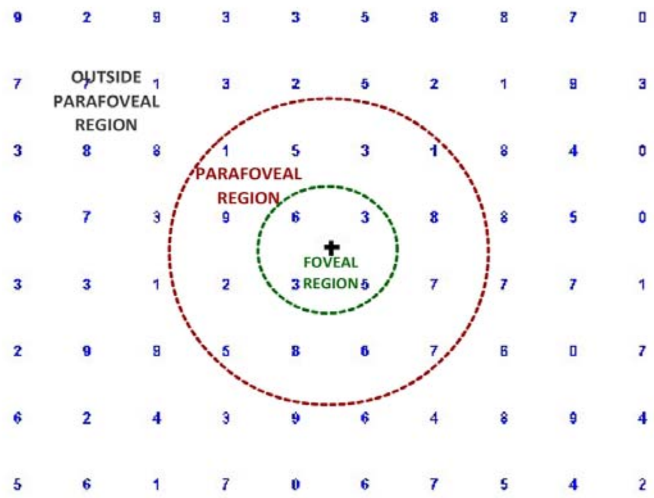


Figure 5-1: Screenshot Experiment 1 single-layer condition

5.2.1.5. Experimental Design

Randomly, one of the digits would change into another digit; for example a 3 into 5, or 1 to 6. Participants were instructed to click on the digit that changed. There were 10 practice trials in which participants received feedback on their performance. There were 148 experimental trials and no feedback was given. Twenty one changes occurred per block. A block was defined based on the transient duration (0, 250, 350 and 450 ms). There were 63 trials for the colour transient condition, 64 for the depth-transient and 21 trials for the no-transient condition.

The change was highlighted by a colour-transient (the changing digit would become red and then revert to blue); or a depth-transient (the changing digit would pop-up to the front layer and then pop-back). The control condition did not highlight the change (No-Transient condition). For the colour and depth conditions, the duration of the transient was 250, 350 or 450 ms. For a short video of this experiment, see Appendix III – Video 1.

Participants had 3500 ms to respond after a change had occurred which was estimated using the KLM method explained in section 5.2.1.1. As mentioned before, in order to account for slower users, a random delay of between 1000 and 2000 milliseconds was inserted between each trial.

5.2.1.6. Procedure

Participants were tested individually. Each participant was given a written description of the experiment along with a set of instructions and an informed consent form (See Appendix IV).

After reviewing the instructions, participants were seated in front of the MLD at a viewing distance of ~55 cm. In order to ensure this viewing distance, the chair was fixed to the floor and a cardboard box was placed in the participants lap. The box served to prevent the participant leaning closer to the screen. No head restraint was used, but participants were asked to keep their gaze in the fixation cross and remain in the same position throughout the experiment. The researcher was present during all trial to make sure that participants follow the instructions. Participants pressed any key to initiate each trial. Then, the 10×8 matrix with a black fixation cross located in the centre of the display appeared in the rear layer of the MLD.

5.2.1.7. Data Analysis

The data was analyzed using several ANOVA tests. A one-way ANOVA within-participants analysis was conducted to evaluate response times and accuracy between the depth transient, the colour transient and the non-transient condition.

Later, a 2×3×3 repeated measures ANOVA within-participants analysis was performed to evaluate colour vs. depth.

Before going any further, there are a few terms and assumptions that are essential to understand. In summary, the term *Repeated Measures* is used when the same participants take part in all conditions of an experiment (Field, 2009). Other tests based on parametric data assume that data points are independent. Within-participant factors like reaction time remain largely consistent throughout the experiment; a participant's reaction time in one condition will not vary independently of their time in other conditions (Field, 2009). The F-test in ANOVA depends upon the assumption that scores in different conditions are independent. When using repeated measures ANOVA, this assumption is violated because scores taken under different experimental conditions are likely to be related because they come from the same participants. Therefore an additional assumption has to be made in order for the F-test to become accurate. This assumption is called the assumption of sphericity (Baguley, 2004; Field, 2009).

The *sphericity assumption* can be thought of as an extension of the homogeneity of variance assumption in independent measures ANOVA. Therefore sphericity relates to the equality of the variances of the differences between the levels of the repeated measures factor (Field, 2009). Sphericity requires that the variances for each set of difference scores are equal. There are two broad approaches to dealing with violations of sphericity. The first is to use a correction to the standard ANOVA tests. The second is to use a different test that does not assume sphericity like MANOVA.

MANOVA tends to have less power than ANOVA and requires rather large sample sizes because the number of cases in each category must be larger than the number of dependent variables (Hill *et al.*, 2005; Tabachnick *et al.*, 1996). In some instances MANOVA cannot be applied specifically when there are few participants in the design and many levels on the repeated measures factor (Hill *et al.*, 2005). For this particular analysis, the researcher decided to use the corrections for repeated measures ANOVA instead of a different test.

The best known corrections are those developed by Greenhouse and Geisser and Huynh and Feldt (Baguley, 2004; Field, 2009; Howell, 2009a). Sphericity is represented by the Greek epsilon (ε). Girden (1992) recommends that when ε is greater than 0.75, then the Huynh-Feldt correction should be used but when ε is lesser than or equal to 0.75, the Greenhouse-Geisser correction should be used instead. In any case, each of these corrections attempts to adjust the degrees of freedom in the ANOVA test in order to produce a more accurate significance (*p*) value. If sphericity is violated the *p* values need to be adjusted upwards and this can be accomplished by adjusting the degrees of freedom downwards (Baguley, 2004; Field, 2009; Mauchly, 1940).

A deeper analysis of sphericity is beyond the scope of this thesis. For a detailed introduction to sphericity, read Chapter 13 from Field 2009. For an explanation on how to calculate the corrections, read Girden 1992. For the calculations on how to assess the severity of departures from sphericity, read Mauchly 1940.

5.2.2. **Results**

Performance was assessed on accuracy based on percentages of correct clicks and response times on milliseconds. For a full set of tables, see Appendix V.

5.2.2.1. Depth, Colour and Non-transient

A one-way ANOVA was conducted to test for an effect on detection among three transient types. Response times across the three transient types were no significant, $F(2, 63) = .414$, $p > .05$. Accuracy, on the other hand, differed significantly across the three transient types, $F(2, 63) = 140.557$, $p < .05$. Tukey post hoc comparisons of the three transient types indicated that depth ($M = .93$, $SD = .07$) was significantly more accurate than colour ($M = .81$, $SD = .07$), $p = .001$ and than the no- transient type ($M = .47$, $SD = .15$), $p < .001$.

5.2.2.2. Depth vs. Colour - Accuracy

Accuracy was assessed as the percentages of correct clicks. A correct click occurred when a participant detected a change and correctly identified it by clicking the digit that changed.

As mentioned before, a 2 (transient types: colour, depth) ×3 (duration: 250,350,450 ms) ×3 (visual regions: foveal, parafoveal, outside the parafoveal region) Repeated Measures ANOVA was conducted. Paired t-test analyses were conducted to compare the colour, depth and non-transient conditions.

There was a significant main effect of the transient type, $F(1,18)=107.78$, $p<.01$. Post Hoc[7] analysis indicated that the depth-transient provoked more accurate responses ($M=.93$, $SD=.065$) than the colour-transient ($M=.81$, $SD=.073$), $t(21) = -8.148$, $p < .0125$.

The main effect of the transients' durations yielded an F ratio of $F(2,36)=7.37$, $p<.01$. The main effect of the visual regions was also significant yielding an F ratio of $F(2,36)=100.36$, $p<.01$. The interaction between the transient type and the transient duration was not significant, $F(2,36)=2.59, p>.05$.

The rest of the interactions violated the assumption of sphericity (see section 5.2.1.7). Thus, the degrees of freedom were corrected using the Greenhouse-Geisser estimates of sphericity ($\varepsilon \approx .53$). The interaction between the transient type and the visual region was significant whether it was corrected for sphericity or not, $F(2,36)=90.13$, $p<.05$. Inspection of figure Figure 5-2 suggests that while within the foveal and parafoveal regions detection of the colour and depth-transients provoked almost 100% accuracy; outside the parafoveal region, participants were significantly more accurate when the change was highlighted with the depth-transient (Figure 5-2). Post hoc analysis confirmed that participants were highly accurate when changes were highlighted in depth outside the parafoveal region ($M=.86$, $SD = .12$) compared to colour outside the parafoveal region ($M=.60$, $SD = .13$), $t(21) = -10.175$, $p<.0125$. Participants' accuracy when the change was highlighted with colour and located in foveal region ($M=.99$, $SD=.021$) was almost the

[7] All post hoc analyses were performed using a Bonferroni correction. For experiment 1, there were four comparisons made, therefore $\alpha = 0.0125$

same as when the change was highlighted with depth in the foveal region (M=.996, SD = .017), t (21) = -.153, p>.05. A similar effect occurred when the change was located in the parafoveal region, the colour transient (M=.986, SD =.025) and the depth transient (M=.983, SD =.028) did not differ significantly, t (21) = .568, p>.05.

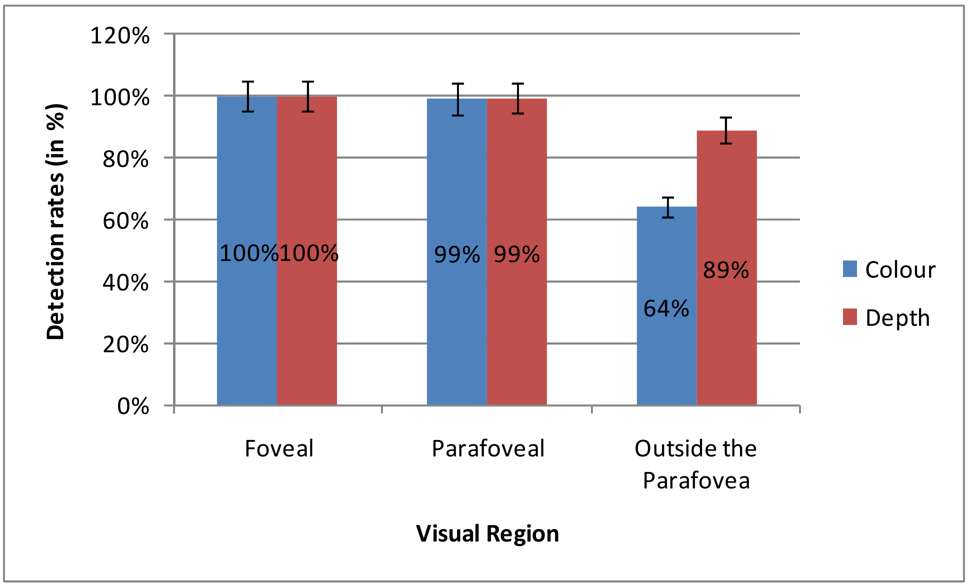


Figure 5-2. Detection rates based on the visual region (Vertical bars depict 95% confidence levels about the mean in each condition)

The interaction between the transient durations and the visual region was also significant, F(2.04,36.75)=5.11, p<.05. The interaction between the transient type, the duration and the visual region was borderline significant yielding an F ratio of F(2.18,39.32)=3.11, p=.051. Closer examination to Figure 5-3 indicates that for those changes presented outside the parafoveal region, the depth-transient provoked the most accurate responses when the cue was presented for longer than 350 ms. However further statistical analysis is required.

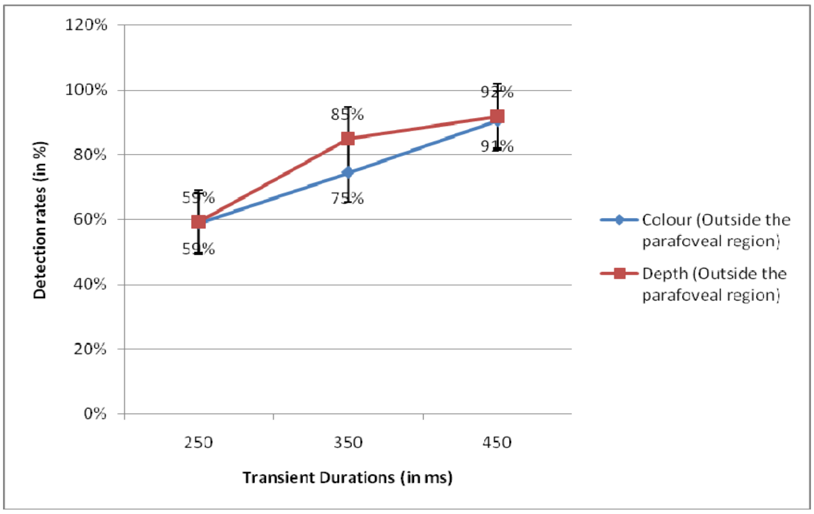


Figure 5-3. Detection rates of changes outside the parafoveal region based on transient duration and transient type (Vertical lines depict 95% confidence levels about the mean in each condition).

Paired t-test analyses showed that the accuracy of participants when changes were not highlighted with any transient decreased significantly. The accuracy of detection for the non-transient condition (M=.44, SD=.145) was significantly lower than to the depth transient (M=.93, SD=.065), t (21) = 19.255, p<.001, and to the colour transient (M=.81, SD=.073), t (21) = 16.118, p<.001.

5.2.2.3. Depth vs. Colour - Response Times

A 2×3×3 Repeated Measures ANOVA including two types of transients (colour, depth), three transients' durations (250,350,450 ms), and three visual regions (foveal, parafoveal, outside the parafoveal region). The main effects and the interactions were non-significant.

Despite the initial hypothesis that changes highlighted with colour would provoke faster response times than those highlighted with a depth-transient, the analysis of response times yielded non-significant results. This suggests that in the shortest duration conditions, only the participants who notice the change quickly, responded. In conditions with longer durations the participants have more time to notice the change, and therefore were able to respond more slowly. Since longer duration conditions allow slower response times, their mean response is slower than short duration conditions. This effect occurred regardless of the type of transient that was presented.

When plotting every data point of response times of correct detections sorted from the smallest to the largest, it became evident that response times for the colour and the depth-transient have almost the same number of correct responses if capped at 1500 ms. Above the 1500 ms response time, the depth-transient presents more correct responses than colour (Figure 5-4).

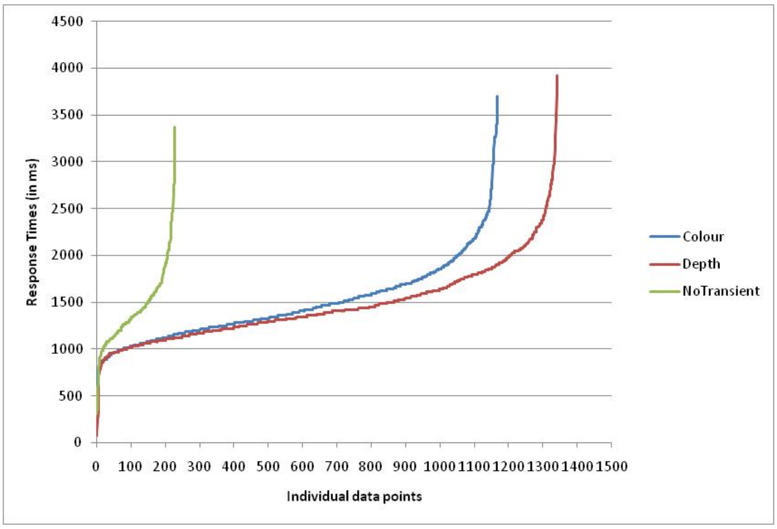


Figure 5-4. Response Times in ms of correct clicks plotted by individual data points and classified by transient type.

5.2.3. Discussion and Limitations

The main purpose of this experiment was to determine whether depth is sufficient as a visual cue. Can observers detect brief changes in the depth of a stimulus? If so, is detection better or worse than a similarly brief change in colour? Going back to the discussion in Chapter 4 (Section 4.4.1), colour is processed in parallel, but the evidence for depth in visual search tasks is inconclusive. Hence, by comparing participants' depth detection performance to the non-transient condition, it was possible to determine whether depth is a sufficient visual cue and by comparing it to colour, it was possible to estimate its efficacy relative to one of the most widely visual cues.

This experiment evaluated the effect of MLD's depth as a transient to highlight expected changes. Contrary to our initial hypothesis, that the colour-transient would provoke faster response times than the depth-transient, the effect was non-significant. In fact, for the entire experiment, longer transient durations lead to slower response times. This effect could be explained in two ways:

In the first place, it is possible that, in the shortest duration conditions, only the participants who noticed the change quickly responded, while in conditions with longer durations the participants had more time to notice the change, and were, therefore, able to respond more slowly. Since longer duration conditions allow slower response times, their mean response is slower than short duration conditions. So on average, there was no effect of duration, type of transient and visual region on response times.

Second, a more likely explanation that stems from the transient durations is that each transient produced an onset (an abrupt flash) when the digit changed in depth, and an offset when the digit went back to its original state. These two "flashes" might have increased the likelihood of detection since participants had two chances to detect the change. The effect of the onset and offset is speculative and further research is required. Nevertheless, when evaluating the correct response times sorted from the fastest to the slowest, the results suggest that the depth-transient could be detected as fast as colour.

The analysis of detection rates suggested that changes presented in the visual periphery, or what was referred as *outside the parafoveal region*, were detected more accurately if highlighted with a depth-transient. Therefore, changes presented at distances greater than five degrees of visual angle and highlighted with a depth-transient are detected with greater accuracy than those changes highlighted with colour.

The significant differences in accuracy showed that depth-transient had a powerful effect. The lack of significant differences in response times is an indicator that the depth-transient could be as good as colour for directing attention to expected changes. Despite these positive results, the exploratory nature of this experiment presented some limitations:

a) The current design is constrained by the lack of one condition: a motion-transient. The fact that the changed digit pop up to the front layer meant that a motion signal was produced. The higher accuracy obtained outside the parafoveal region might have been due to an effect of motion rather than depth. If we recall chapter three, the human eye has more rods than cones. The rods are better motion sensors, and therefore the human eye can detect motion better with its peripheral vision, since it is primarily rod vision. Adding this condition would have allowed to partitioned out the possibility of a motion effect.

b) The latency in response times could have been affected because the stimuli were presented on the rear layer which has the disadvantage of slightly blurring the images. However, the effect of depth couldn't be evaluated if the stimuli were presented on the front layer because the pop-up effect would have been eliminated.

5.3. Experiment 2: The effect of depth on the detection of unexpected events

This experiment evaluates the effect of the depth of the MLD in the detection of an unexpected event. The experimental design is similar to that of Most *et al*. (2000). Several studies, including that of Most *et al* (2000), have replicated Mack and Rock's (1999) technique to analyse the detection of unexpected events (See section 4.2). This technique guarantees that the participant would neither be expecting nor looking for the object of interest. It permits only one-true critical inattention trial per subject because for subsequent trials, participants are likely to be expecting the critical stimulus to appear. Therefore, a large number of participants are required.

Results from Experiment 1 indicated that the depth of the MLD provokes highly accurate detection rates. Based on those results, it was assumed that an unexpected event presented on the front layer would be more noticeable than one presented on the rear layer.

The effect of eccentricity on target detection was also analysed expecting that detection of the unexpected event would be inversely proportional to its eccentricity from the fixation point, as has been shown in previous studies (See Chapter 4, and Carrasco *et al.*, 1995; Most *et al.*, 2000; Wickens *et al.*, 2003).

5.3.1. **Materials and Methodology**

5.3.1.1. Participants

Sixty participants volunteered to participate without payment. Data from eight participants were excluded because they were considered outliers. Among the outliers, one participant reported losing count of the hits; six participants' total count was more than two standard deviations away from the mean of the other participants in that condition, and another participant did not see the cross in any of the trials.

Of the remaining 52 participants, 20 were males and 32 were females with a mean age of 26 years (SD=6) varying from 17 to 51. All participants had normal or corrected to normal vision. All participants took the Ishihara colour test and presented normal colour vision (http://tinyurl.com/26wm3c). Stereoacuity was not measured.

5.3.1.2. Equipment

Stimuli were presented on the same MLD used for the previous experiment. It was set at a resolution of 800×600 per screen.

5.3.1.3. Stimuli

The MLD was placed at approximately 60 cm from the participant. The display contained eight circles with a radius of 12.5 pixels. The circles were equally divided in two groups: blue and red. A black bar was located at 2.6

cm from the bottom of the display. The bar was 1.3 cm wide and 12.8 cm long, and it was presented on both screens (Figure 5-6).

The unexpected object was a black cross with the same radius as the circles. The circles moved in pseudo-random patterns. They would move in straight lines until they bounced against the side of the display or the black bar, in which case their heading would change, "bouncing" in a physically unrealistic direction. The pattern was determined by a series of scripts, and all participants saw the same display patterns in the same order.

5.3.1.4. Experimental Design

Participants were divided in two groups: those counting the blue circles and those counting the red circles. For the single-layer condition (SLD), all stimuli were presented on the back layer of the MLD. For the double-layer condition, the circles were equally distributed in both layers, but the unexpected event was presented on the front or the back layer (Figure 5-5).

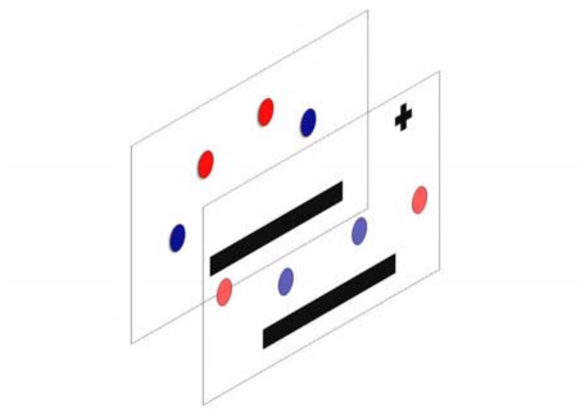


Figure 5-5: Schematic representation Experiment 2 MLD condition

The unexpected object's position was also manipulated so it will appear at two locations: for the *near* condition, the cross was located at 6.58cm from the black bar; and for the *far* condition, the cross was placed at 13.68 cm from the black bar. Figure 5-6 shows a schematic representation of the experiment. The dotted gray lines are approximate estimations of the visual angle measured from the black bar. The green lines, arrows, measurements and dotted grey lines were not presented during the experiment. For a short video of this experiment, see Appendix III – Video 2.

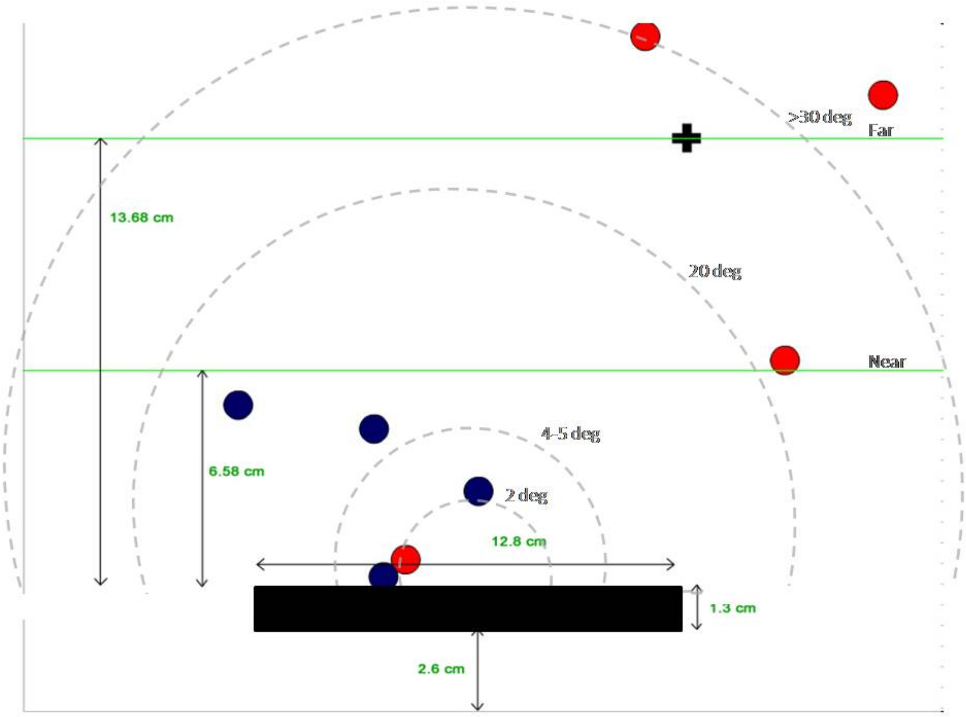


Figure 5-6. Schematic representation Experiment 2 single-layer condition

5.3.1.5. Procedure

All trials were completed in one session which lasted five minutes. Participants were given a set of instructions and an informed consent form (See Appendix IV). Participants were seated approximately 55 cm from the screen. As in Experiment 1, a fixed chair and cardboard box were used to maintain this viewing distance. No head restraint was used, but participants were asked to keep their heads straight and remain in the same position throughout the experiment.

Observers were instructed to silently count the number of times the circles of a designated colour hit the black bar. No mention was made of unexpected events.

Participants had five trials and participated only in one condition. The first two trials were "standard" trials, in which no unexpected events were presented. The third trial was the critical inattention one which tested detection when participants were totally unaware of the possible presence of another object. Fifteen seconds into the third trial, a black cross entered from the right side of the display, moved horizontally in a linear path across the screen, and exited the left side of the display. The cross was visible for a total of five seconds.

After completing the third trial, participants were asked to report whether or not they had seen anything other than the circles, and if they had seen something else, to describe it.

Observers then completed a fourth trial on which the cross appeared again travelling on the same path. For this trial, participants were implicitly aware of the existence of another object. After completing this trial, participants answered the same questions.

In the fifth trial, observers who did not see the cross in the last two trials were instructed to watch the display without counting the number of times the circles hit the black bar. Having been alerted by the previous questions, observers were now aware that another object could appear. Furthermore, their attention was not otherwise engaged, so this trial tested perception without a monitoring task. For the rest of the participants the task remained the same. After this trial, observers completed a questionnaire identical to the first two.

5.3.1.6. Data Analysis

Data was analyzed using Fisher's Exact Test. The Fisher's Exact Test is a way of computing the exact probability of a chi-squared statistic. The chi-square test assumes that the sampling distribution of the test statistic has an approximate chi-squared distribution (Field, 2009). The larger the sample is, the better the approximation becomes (Hill *et al.*, 2005). When the expected frequencies are too low, it is safe to assume that "the sample size is too small and that most probably, the sampling distribution is too deviant from a chi-squared distribution" (Field, 2009). This experiment presented some categorical variables with one or more expected frequencies being less than or equal to five. Therefore, the Fisher's Exact Test was used instead of the chi-square because it gives an exact *p* value and works fine with small sample sizes.

Cramer's V is a way of calculating correlation in tables which have more than two rows and two columns which was the case for this experiment where the contingency tables were 2×2×3. The Chi-square and the Fisher's Exact Test indicate that there is a significant relationship between variables, but these tests do not indicate how significant and important the relationship is (Changingminds.org, 2010). Cramer's V is a post-test to give this additional information. It is a measure of strength of association between two categorical variables, used when one of these variables has more than two categories (Field, 2009). V is calculated by first calculating chi-square, then using the following calculation: V = SQRT(χ^2 / (n (k - 1))), where χ^2 is chi-square and k is the number of rows or columns in the table. Thus, the following analysis used a Fisher's Exact Test and the strength of association between categorical variables was measured using Cramer's V.

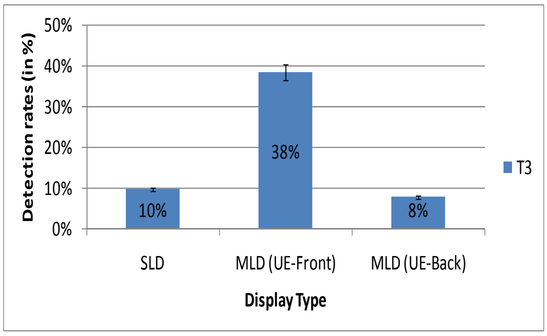
5.3.2. **Results**

Participant performance was assessed according to the detection rates of the unexpected event. Accuracy in the counting task was measured by comparing the reported number of times the circles hit the bar with the actual number. For a full report with all contingency tables, see Appendix VI.

5.3.2.1. Detection

A Fisher's Exact Test was performed to examine whether detection of the unexpected event was equal among the single-layer condition, the MLD condition with the unexpected event presented in the front layer, and the MLD condition with the unexpected event presented in the back layer. The detection rates were significantly different for the third, p=.013, and the fourth trial, p=.048. The results indicate that when the unexpected event is presented on the front layer and participants are not expecting it, the unexpected event is detected 38% of the time; but if its on the rear layer is detected only 8% of the time. The association between the layer in which the unexpected event is presented and its detection is strong (*V*=.40). For the fourth trial, when participants are implicitly alerted about the existence of another object, the detection rates increased to 42% when it appeared on the front layer, while participants only detected the unexpected event 12% of the time when it traversed through the back layer (Figure 5-7).

(a) 3rd Trial **(b) 4th Trial**



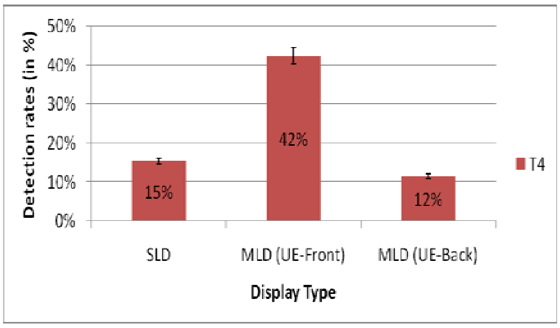


Figure 5-7. Detection rates for the third and fourth trial (Vertical bars depict the 95% confidence intervals about the means in each condition)

An analysis of eccentricity (Figure 5-8) indicated that there was a significant association between the position of the unexpected event and its detection during the third trial, p=.011, with a fairly strong effect (*V*=.38). The association was non-significant for the fourth trial, p=.232.

(a) 3rd Trial ($p<.05$) **(b) 4th Trial ($p>.05$)**

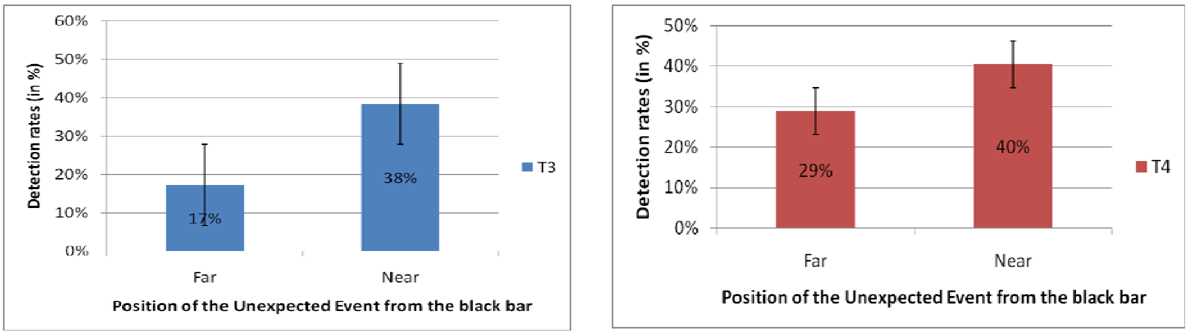


Figure 5-8. Overall detection rates based on the eccentricity of the unexpected event for the (a) third and (b) fourth trial (Vertical bars depict the 95% confidence intervals about the means in each condition).

A Fisher's Exact Test was performed on a 2×2×3 contingency table evaluating the association between two positions of the unexpected event (Near, Far), two detection categories (yes, no) and three display types (SLD, MLD–UE Front, MLD–UE Back). There was a non-significant association between the three variables.

The luminance of the circles seemed to have influenced the detection of the cross. There was a significant difference in the detection of the unexpected event for the third, $p=.024$ and fourth trial, $p=.007$ indicating that participants detected the unexpected event more often when monitoring the blue circles compared to the red ones (Figure 5-9).

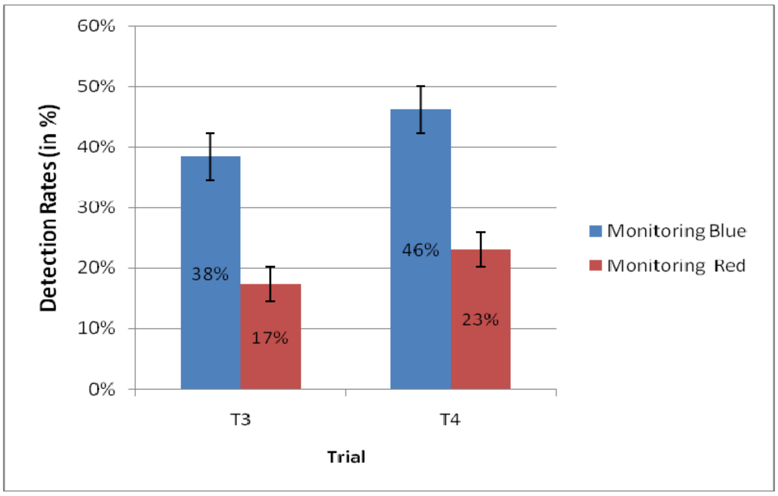


Figure 5-9. Detection of the unexpected event when monitoring the blue and the red circles (Vertical bars depict the 95% confidence intervals about the means in each condition).

A Fisher's Exact Test was performed on a 2×2×3 contingency table evaluating the association between two colours of the monitored circles (Blue, Red), two detection categories (yes, no) and three display types (SLD, MLD–UE Front, MLD–UE Back). There was a non-significant association between the three variables for the third trial, but the similar luminance between the blue circles and the black cross provoked a significance association in detection rates for the fourth trial when the unexpected event was presented on the front layer, p=.006 (Figure 5-10).

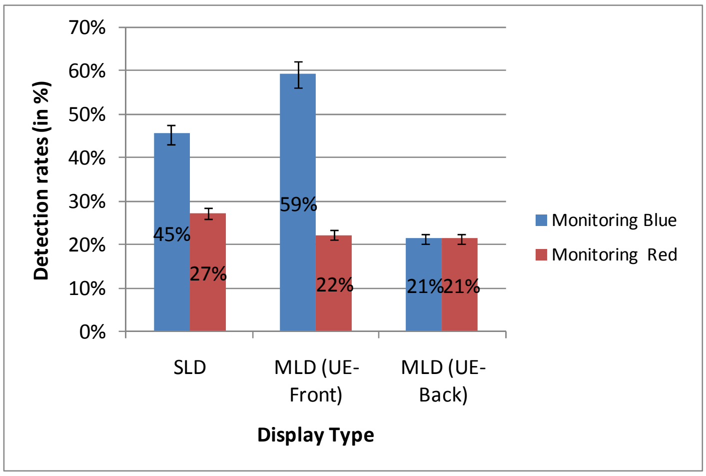


Figure 5-10. Detection rates for the fourth trial - unexpected event front layer (Vertical bars depict the 95% confidence intervals about the means in each condition).

5.3.3. **Accuracy**

Accuracy analysis was based on the precision of the count. Trial 3 presented 19 hits and trial 4, 20 hits. Analysis of the standard deviation indicates that participants were more accurate in the first two trials (M_1=18, SD_1=0.9; M_2=19, SD_2=0.96) than in the third and the fourth (M_3=19, SD_3=1.47; M_4=20, SD_4=1.74).

A paired t-test analysis found a significant decrease in accuracy for those who detected the cross in the third trial compared to the first and second; surprisingly there was no significant difference in the reduction of accuracy for those who did not detected the cross ($p_{detection} < .05$; $p_{non\text{-}detection} > .05$) .

An analysis of the mean of the error showed that participants were more accurate in counting the circles when they were distributed in two layers compared to a single layer (Table 5-1).

Table 5-1. Mean error for Trials 1 and 2 vs. Trial 3

	Mean Error T1 &	Mean Error T3
Circles in rear layer[8]	0.74	1.24
Circles distributed in 2 layers	0.33	0.56

5.3.4. **Discussion and Limitations**

The results demonstrated an important effect of depth and eccentricity on the detection of an unexpected event, but failed to show a significant relation between these variables. This suggests that an unexpected event becomes more noticeable if: (1) it is on the front layer, or (2) it occurs close to the focus of attention.

Consistent with the results of previous experiments on Inattentional Blindness, detection of the unexpected event in the fourth trial increased compared to the third trial. This rise is due to the fact that participants are implicitly aware that another object might appear on the screen (See section 4.2 and reviews Mack, 2007; Simons, 2007).

Detection of the unexpected event also increased for those participants monitoring the blue circles. It has been shown in previous research that similar luminance or shape of the unexpected object to the attended items had greatly affected the likelihood that people would notice the unexpected object (Most *et al.*, 2005; Most *et al.*, 2001; Simons *et al.*, 1999). Their results demonstrated that when observers are engaged in a challenging task that requires selective processing, they establish an attentional set on the basis of the dimension critical to proper selection.

[8] The smaller the number, the higher the accuracy

When the unexpected object matched the preset characteristics of the attentional set -the participant was monitoring the dark blue circles- then, the participant was more likely to notice the black cross. However, detection of the black cross decreased when monitoring the red circles because, in that case, the black cross did not match the participant's attentional set. It is important to note that this effect was only significant during the fourth trial when the unexpected event was presented on the front layer.

Unfortunately further comparisons with Most's study are not possible because the distances used for the near and far conditions in Experiment 2 are much larger than the distances used by Most *et. al* (2000). While they set the near condition at 2.4 cm and the very far condition at 5.9 cm from the line that was the focus of attention, Experiment 2 was set up so that the unexpected event in the near condition appeared at 6.58 cm from the bar.

Accuracy in counting the hits increased when the stimuli were distributed in two layers. Previous studies on Inattentional Blindness have not reported results on accuracy of the main task; therefore, comparisons were not possible to draw.

This experiment found a significant effect of depth in the detection of an unexpected event. Nevertheless, it is important to clarify that this was an exploratory study and suffered from the lack of one condition: a single layer condition in which all the stimuli were placed on the front layer. Placing all the stimuli on the front layer could have eliminated the possibility that the difficulty in detecting the cross was due to the slight blur that the rear layer presents. However, both MLD screens affect the quality of the image. While the rear layer is slightly blurred by the Perspex layer, the front screen makes images translucent. Additionally, if this condition was added, that would have meant that the sample size had to be increase by at least 60% taking into account that in any Inattentional Blindness experiment, observers can only participate once.

Regarding the sample size, a sample of 52 is rather small. Given the fact that observers could only participate in one condition and they could only receive one critical inattention trial, about five data points were left in some conditions. This sample size might have affected the power of the experiment, and although the main effects presented strong levels of association, the relation between them provided inconclusive results.

Overall, the detection of the unexpected event was five times higher when it was presented on the front layer and stimuli were distributed in both layers compared to the single layer presentation. When stimuli were distributed in two layers, the detection of the unexpected event was four times higher when it was presented on the front layer compared to the back layer.

5.4. Conclusions

The complexity of many systems due to their high volume of rapidly changing information has raised concerns about change detection in operational environments. The nature of the funding suggested that the MLD could be used to enhance operators' change detection if used as an alerting tool.

The objective of these experiments was to determine whether the MLD's depth had an effect on detection of expected and unexpected events when used as an alerting tool. Previous research on visual search has shown that colour is a very effective cue to guide attention because it is processed preattentively. Results from studies that have analysed various depth cues have been inconclusive and it still remains unanswered whether depth in general or a specific type of depth cue is processed in parallel.

Although results from Experiments 1 and 2 suggest an effect of the MLD's depth on change detection, it is not yet recommended for using the MLD commercially or operationally. The design of the experiments detailed in this chapter had some limitations. Some of these limitations are because of the restrictions presented by the MLD, while others are due to the limitations of the experimental design.

The quality of the images on both layers of the MLD is affected by either the backlight source or the Perspex layer, making images translucent on the front screen, and producing a slight blur on the rear screen respectively. Some have argued that this slight reduction on the quality of the images could have affected perception and detection. However, previous research had shown that the use of translucency and alpha-blending does not affect perception and that even when using full colour images or videos, detection is impaired if our attention is diverted to a different task. Remember Neisser and Becklen (1975) experiments. While investigating the mechanism of "selective looking" (See chapter 4), they asked participants to look at two optically superimposed video screens, on which two different kinds of things were happening. Participants were required to follow the action in one "screen" and ignore the other. Results showed that participants could easily deploy their attention to one of the videos, although both were present in the same fully overlapped visual field. Events in the unwatched screen were rarely noticed. Most importantly, about 50% of participants were able to detect unexpected events in the video they were watching (Neisser, 1976; Neisser *et al.*, 1975).

Many cognitive psychologists found these results interesting but were somewhat less convinced of the importance of the failures to notice unexpected events because the video superimposition gave the videos an odd appearance that is not typically experienced in the real world (Simons *et al.*, 1999). These counterarguments were eliminated by Simons and Chabris who replicated Neisser's experiments without the video superimposition. By showing a full-colour video of one event (the basketball game), their results showed that about 50% of participants detected the unexpected event. Therefore, inattentional blindness occurred at similar rates even without the video superimposition (Simons *et al.*, 1999).

The superimposition used by Neisser is similar to the one produced by the MLD. Images on the front layer become translucent which is comparable to using alpha blending and images are superimposed when information is presented on both layers. In Experiment 2, inattentional blindness occurred as expected but the results suggest that depth had a positive effect on detection. Although further research and replication is required, the superimposition and the translucency of the front layer did not affect perception. Whether stimuli are distributed on two depth planes or one, our cognitive limitations come into play when our attention has to be diverted to a different specific task.

Another technological restriction of the display technology is its binocular disparity. The MLD produces a binocular disparity of 8 arc seconds, but, as described in chapter 4, De la Rosa (2008) suggested that in order to process depth cues in parallel the binocular disparity of the display should be at least 6 arc minutes for visual search tasks (De la Rosa *et al.*, 2008). This mean that although participants are able to deploy their attention in depth and perceive changes highlighted with depth, the depth of the MLD cannot be processed in parallel and this might affect respondents' response times.

The limitations of the experimental design means that we cannot rule out that the high rates of detection in Experiment 1 could have been due to radial motion as opposed to depth. It also means that in Experiment 2, although it seems unlikely, the inattentional blindness rates in the control condition might have been affected by locating all the stimuli on the rear layer.

In any case, Experiment 1 indicated that depth as a visual cue not only provoked very accurate responses, but an analysis of response times also suggested that the depth-transient could be as good as colour to guide attention to an expected change. The analysis of eccentricity yielded interesting results. Within the foveal and the parafoveal region, colour and depth-transients resulted in almost 100% accuracy. However, changes located outside the parafoveal region were detected more accurately when highlighted with a depth-transient.

Experiment 2 showed that an unexpected event can be detected more often if it is presented on front layer or closer to the focus of attention or presents similar features to the monitored object. Experiment 2 showed that participants could detect the unexpected event with high accuracy when it was presented on the front layer. When compared to a single-layer presentation of the stimuli, participants detected the unexpected event five times more often when it was on the front layer. When stimuli were distributed in both layers, the detection of the unexpected event was four times higher when the cross was presented on the front layer than when it presented on the back layer.

Consistent with other studies, spatial proximity of the unexpected event to the focus of attention played an important role in detection. During the inattention trial, participants were twice as likely to detect the unexpected object when it was closer to the focus of attention (the black bar) than when it was located in the far upper region of the display.

5.4.1. **On the effect of onsets and offsets**

An interesting point for discussion that will need further research is the effect of the transient duration that was mentioned in section 5.2.3. Previous research has evaluated abrupt onsets and while some argue that abrupt onsets capture attention only if the observers are set to look for them (Mulckhuyse *et al.*, 2008), results seem to indicate that abrupt onsets are processed preattentively (Yantis *et al.*, 1984). The effect of the onset in Experiment 1 might not have been processed preattentively but it might have had an effect on detection. By manipulating the transient durations, it meant that there was an onset when the digit changed and an offset when the digit returned to its original state. It might have been the case that participants detected those changes highlighted in depth with such a high accuracy because they had two abrupt flashes: one at onset and one at offset. The longer latency can also be explained by this because the detection of the change could have occurred either at onset or at offset, giving the participants more time to react. Whether the change of colour had the same effect is hard to conclude, but further research is required to establish whether the onset/offset of the depth-transient captured attention increasing accuracy.

5.4.2. **On MLDs, SLDs and LCDs**

Another point that seems important to note is the difference between single-layered LCD monitors and the multi-layered LCD monitors in order to be aware of the implications that the display technology had on the design of the experiments.

As we know, the MLD used for these experiments comprises two LCD displays. The MLD needs a backlight source because LCDs do not produce light themselves unlike cathode ray tube (CRT) displays. The backlight is used to produce light in a manner similar to a CRT display (Wiley, 2008). LCD monitors need a mechanism to regulate the light intensity of the screen's pixels. The most common element is a polarizing filter to polarize the light from the source in one of two transverse directions and then passing it through a switching polarizing filter to block the path of undesirable light (Wiley, 2008). LCD monitors reproduce colours by manipulating light waves and subtracting colours from white light.

These characteristics become limitations when the screens are stacked together in the MLD. First, the need for a backlight source makes the MLD technology hard to use in portable devices because the need of a powerful battery to produce enough light for two displays. Second, the mechanism to regulate light intensity makes it harder to maintain colour accuracy and vibrancy whereas in other types of displays like CRT or OLED[9] monitors, colours are more stable and easier to manipulate.

Designing experiments to evaluating depth in general is hard, but the MLD makes it harder because its depth cannot be manipulated. Not only because the physical separation between the screens is fixed, but because the depth of the MLD is binary. The MLD's software does not allow to present different levels of depth between the screens. An image can be either at the back or at the front but not in between. So although autostereoscopic depth is what

[9] Colour is richer and more realistic. It is generated by organic phosphorus using active matrix technology.

makes the MLD special, it is also a drawback since different layers of depth can be easily simulated on a single-layered screen but not so easily produced by the MLD.

These characteristics limit the use of the MLD. However, these characteristics could probably be used on its advantage. For instance, on Experiment 2, when the cross is presented on the front layer, it means that it was always on sight. The cross becomes translucent when another object is passing behind it on the rear layer, but it is never occluded by the rest of the stimuli. This main difference compared to single-layered displays may be one of the reasons for finding such a positive effect on the detection of the unexpected event on the front layer. Additionally, the experimental design used for Experiment 1 was specifically devised for the MLD. Although radial motion could be simulated on a single-layered display, depth itself would be more difficult. It was the physical separation of the MLD's screens that made it possible to evaluate the sufficiency of depth as a visual cue.

In summary, despite the limitations presented by the design of these two experiments, a positive effect of the depth of the MLD was found. The MLD as a display technology presents some limitations as well but if used appropriately, it can be useful under certain conditions, such as instances where binary information needs to be highlighted. The results of Experimetnns 1 and 2 suggest that the MLD has the potential to be used as an alerting tool and depth could be used as an alternative visual cue to colour and flashing.

Chapter 6

EVALUATING THE MLD AS A COMPARISON TOOL

Chapter 5 determined that depth can be used as an alerting tool. However, when using alarms and alerts, the system requires detailed models of the operations in order to highlight important changes of information to the operator. This chapter evaluates the effect of the MLD as a comparison tool. Instead of depending on an alerting system to mediate detection, the technology could allow operators to detect differences between two images directly. This chapter describes two laboratory experiments conducted to evaluate the effect of the depth of the MLD to allow comparison of simple and complex stimuli.

6.1. Introduction

The use of alerts, visual, audible or even tactile, has been a traditional solution to deficits in change detection (Obermayer *et al.*, 1999; Wickens *et al.*, 2008). Nonetheless, in complex operational situations, the use of alerts requires detailed models of the operational environment. For instance, in air traffic control, to alert pilots about potential conflicts, the system should be able to project trajectories and predict behaviour based on current trajectory, final destination and current and projected speed of the aircraft.

Rather than building complicated models of the world and relying on the computer to detect those changes so it can alert the operator, it was hypothesised that a system that allows the comparison of two separated states would shift the task of change detection from the realm of memory and cognition to that of perception.

One way to present a pair of states or images for comparison is side-by-side but any kind of coplanar separation will necessitate saccades on the part of the user. These saccades are known to pose a problem for the visual system and cause problems for change detection, because they generate rapid, large-field motion on the retina and change the relationship between the object position in external space and the image position on the retina. The brain must ignore the one and compensate for the other. These reduction of visual sensitivity is known saccadic omission or saccadic suppression (Dornhoefer *et al.*, 2002; Irwin *et al.*, 1998; Ross *et al.*, 2001). The eyes make rapid, saccadic movements from point to point in space several times each second. It is assumed that some mental representation of the visual environment is maintained across eye movements in a trans-saccadic memory, however, the information derived from a single fixation is usually insufficient to base comprehension (Irwin *et al.*, 1998; Parkin, 2000). Studies on change blindness have shown that if the change occurs during a saccade, change detection is highly impaired (Dornhoefer *et al.*, 2002).

The MLD provides a possible solution by allowing the presentation of two images that are spatially separated in depth. In this sense, the two images or states are coaxial (concentric superimposed images), rather than coplanar (lying on the same plane). This coaxial presentation enabled by the layers of the MLD may allow users to compare two states more effectively than the coplanar presentation provided by normal displays, thus facilitating the detection of differences.

To evaluate this hypothesis, two “spot-the-difference” experiments were set up to determine whether participants are able to detect differences in images presented in the MLD.

6.2. The pilot study – Experiment 3 and 4

The pilot study was run in two sessions. The same five participants volunteered for both sessions. In the first session, participants were shown 30 pairs of images that included random shapes. The shapes were presented side-by-side on the front layer of the MLD, or in both layers, centred and superimposed. Each pair could present one of eight manipulations: an item was added, deleted, or displaced, the hue or shade of an item was modified, an item was either bigger or smaller, or an item presented a hole in the middle. Participants were asked to click on the difference when detected. At the end of the session, they were asked to rate the level of difficulty and to suggest other examples of images that they would typically use on a day-to-day basis that might need to be compared. They rated the differences in hue and luminance as the most difficult, and the changes in size and shape as the easiest. Most of them suggested the use of images such as Excel graphs, photographs, cartoons and maps.

Based on these suggestions, pairs of images of Excel graphs, photographs and cartoons were added to the experiment. These images were only manipulated in three ways: an item was added to the image, an item was removed or an item was moved from its original position.

At the end of the session participants were asked to rate the level of difficulty per stimuli and per condition. Differences in shapes and Excel graphs were rated as the simplest to detect, differences in cartoons were slightly more complicated to find, and photographs were the most difficult stimuli.

Response times and correct clicks were analyzed in both sessions. Their response times and accuracy reflected their opinions: On average, they took longer to detect differences in photographs than they did for cartoons, shapes and Excel graphs. The detection was slightly faster for the MLD than the side-by-side, and the deletion of an item was detected faster than its addition, which is consistent with previous studies mentioned in the literature (Rensink, 2002).

Thus, the first experiment evaluated shapes only and included six types of differences. The second experiment introduced photographs, cartoons and Excel graphs to evaluate the detection of differences with more complex stimuli but only evaluated three types of differences.

6.3. General Method

Experiments 3 and 4 used variants of the same "spot-the-difference" method.

6.3.1. Materials and Methodology

6.3.1.1. Participants

A total of twenty four participants volunteered for both experiments. There were five males and 19 females with a mean age of 33 years (SD=10). Participants received a voucher from a local store. All participants had normal or corrected to normal vision. Each observer was screened for normal colour vision, and they all passed the Ishihara test for colour blindness. Observers were not tested for stereoacuity.

6.3.1.2. Equipment

The same 17" MLD as in previous experiments was used but set at a resolution of 1280×1024 per screen.

6.3.1.3. Experimental Design

Images were 600×480 pixels. There were two main conditions: for the single-layer condition (SLD), images were presented side by side with 20 pixels between the images. Both images were presented on the front layer of the MLD. For the MLD condition, the original image was presented on the back layer, and the modified image was presented on the front layer. The images were superimposed and were centred in each screen.

Participants were randomly divided into two groups: A and B. Group A saw the SLD condition first and then the MLD; while Group B saw the MLD condition first and then the SLD. Each group saw different pairs of images for each main condition in order to eliminate any memory effects (i.e. the pairs of images shown to Group A in the SLD condition were the same as those shown to Group B in the MLD condition).

6.3.1.4. Procedure

Participants were tested individually. Each participant was given a written description of the experiment along with a set of instructions (See Appendix VII). After review of the instructions, participants sat in front of the MLD at a viewing distance of approximately 50 cm. In order to ensure this viewing distance, the chair was fixed to the floor and a cardboard box was placed in the participants lap. The box restricted the participant from getting closer to the screen. No chin rest or restraint was used. Half of the participants were allowed to move their heads if needed during the MLD condition but the other half was asked to keep their heads still during the experiment.

The experiments were programmed in Superlab 4.0.7. Response clicks were recorded by Superlab and incorrect answers were tracked on an Excel spreadsheet.

Participants initiated each block of trials and were told to spot the difference between pairs of images. There was always only one difference present for each pair. Participants had to click on the difference in any image. The mouse was not restricted to any layer. During the experimental trials, the images timed out after 20 seconds. Participants were instructed to right-click only if they were unable to distinguish the images.

6.3.1.5. Data Analysis

Response times and accuracy were analysed using Repeated Measures ANOVA within-subjects analysis. For a description of the term Repeated Measures and the underlying assumptions behind the analysis, see section 5.2.1.7.

6.4. Experiment 3: Comparing simple stimuli

This experiment evaluated simple images and introduced six differences to be detected. Participants had a total of 10 practice trials and 52 experimental trials whereas the session lasted an average of 15 minutes.

Four hypotheses were tested: (1) hue differences on the MLD would be detected slower and less accurately than on the SLD; (2) luminance differences in the MLD would not be noticeable provoking a high rate of no-responses; (3) holes differences presented on the MLD would provoke slower and less accurate responses than in the SLD; and (4) changes in size (grow and shrink) would provoke faster response times when shown on the MLD compared to the side by side presentation.

6.4.1.1. Stimuli

Fifty two pairs of images were randomly ordered using a latin-square design. Figure 6-1 shows a sample of the stimuli. Red squares indicate where the differences are. The red squares were not presented during the experiment.

The pairs of images presented six manipulations: (1) changes in *luminance;* (2) changes in *hue;* changes in size so that (3) an item was bigger *(grow),* or (4) an item was smaller *(shrink)* than the original*;* and changes in shape in which an item presented a hole in the middle of the shape (5) *hole-added,* or (6) *hole-filled*.

The *"hole"* condition was divided in two categories to evaluate if they were distinguishable when presented in the MLD: The '*hole-added*' condition presented the shape with the hole on the front layer of the MLD. For the '*hole-filled*' condition, the shape with the hole was on the rear layer. In order to counterbalance this presentation in the SLD trials, the '*hole-added*' occurred when the shape with the hole was on the right-hand side; and the '*hole-filled*', if the shape with the hole was presented on the left-hand side of the screen. For a short video of this experiment, see Appendix III – Video 3.

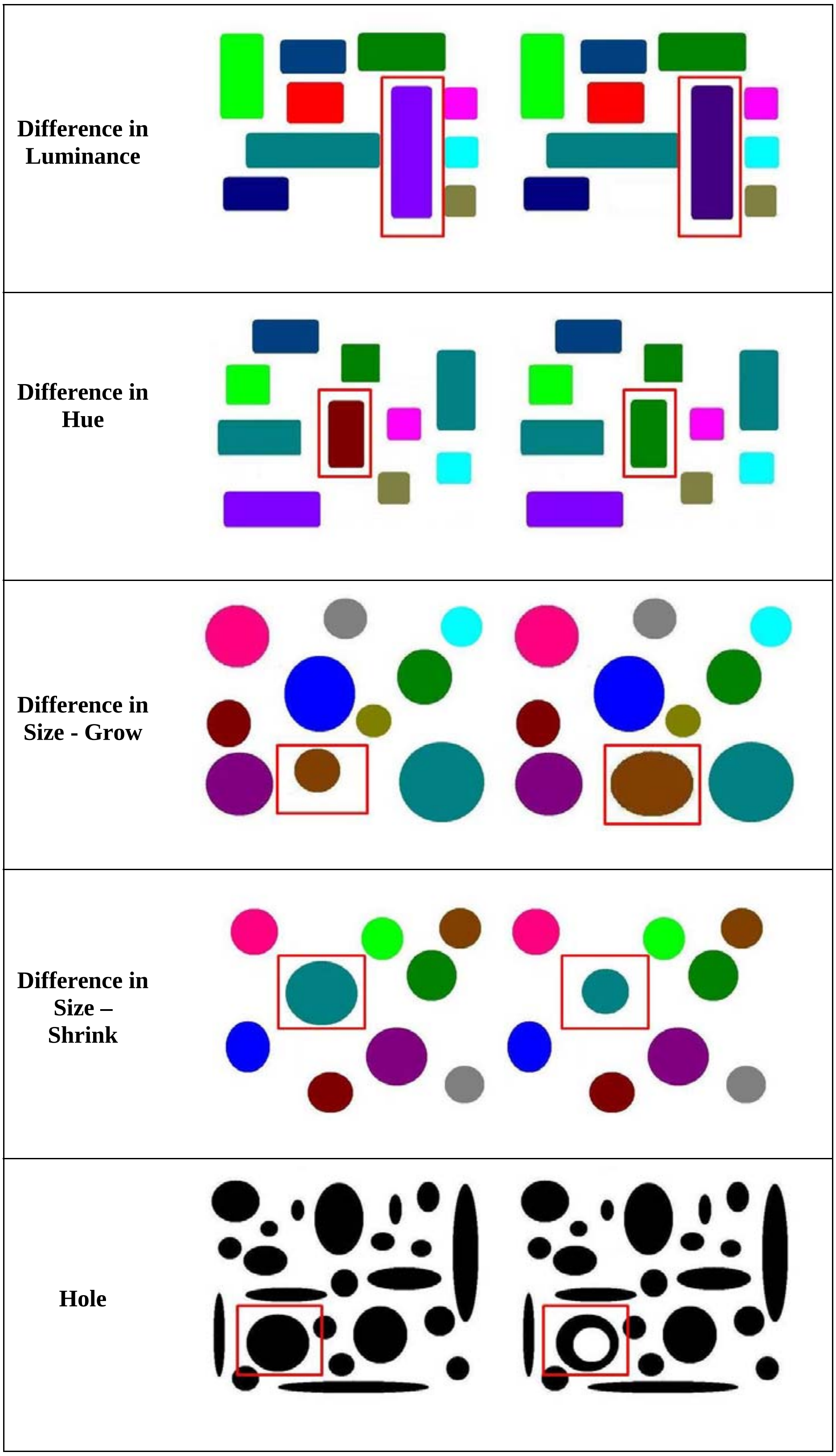


Figure 6-1. Experiment 3 – manipulations shown in the SLD condition

6.4.2. **Results**

A 2×6 Repeated Measures ANOVA was conducted which included two display types (SLD, MLD) and six manipulation types (luminance, hue, grow, shrink, hole-added, hole-filled). Response time was measured in milliseconds and based on correct responses only and accuracy was assessed on correct clicks. For complete report on the tables from SPSS, see Appendix VIII.

6.4.2.1. Response Times

The main effect of the display type was significant, $F(1,17)=14.18$, $p<.01$. The main effect of the manipulation type violated the assumption of sphericity (see section 5.2.1.7) but the results were significant whether the degrees of freedom were corrected or not, $F(5, 85)= 33.87$, $p<.001$. Inspecting Figure 6-2 suggests that participants took almost five seconds longer to detect luminance differences than to detect holes differences but further analysis is required.

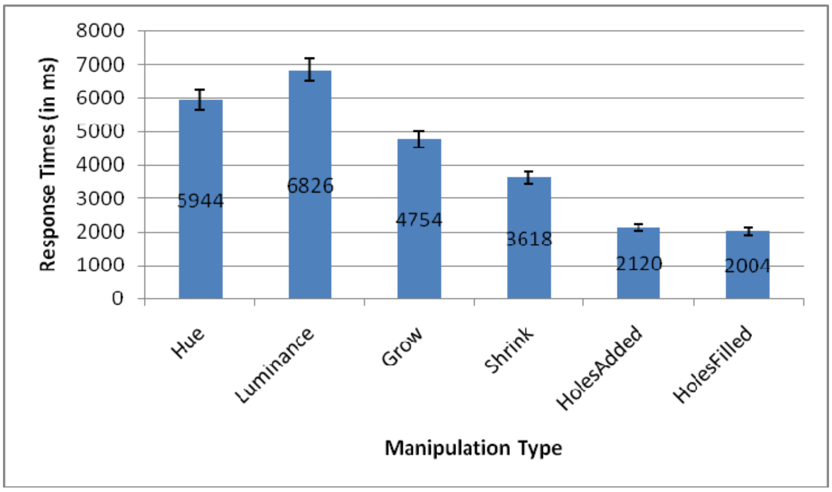


Figure 6-2. Response Times per manipulation type (Vertical bars depict 95% confidence levels about the mean in each condition).

The interaction effect between the display and the manipulation type also violated the sphericity assumption but results were significant whether they were corrected or not. The interaction yielded an F ratio of $F(5, 85)=11.78$, $p<.01$.

Post hoc[10] analyses showed that participants detected differences of size on the MLD significantly faster ($M_{MLDGrow}$= 2838 ms, $SD_{MLDGrow}$=1643 ms; $M_{MLDShrink}$= 1799 ms, $SD_{MLDShrink}$=563 ms) than on the SLD ($M_{SLDGrow}$= 6372 ms, $SD_{SLDGrow}$=2295 ms; $M_{SLDShrink}$=5673 ms, $SD_{SLDShrink}$=2419 ms), t_{Grow} (19)= 5.706, $p<.005$ and t_{Shrink} (19)=7.196, $p<.005$.

The detection of differences in hue when presented on the SLD (M_{Hue}=5526 ms, SD_{Hue}=2242 ms) and MLD (M_{Hue}=6340 ms, SD_{Hue}=3781 ms) was not significant, t (19) = -1.073, $p>.005$.

Post Hoc analyses also indicated that there was no significant difference in the response times when detecting holes in the MLD (M_{Hole_add}=2275 ms, SD_{Hole_add}=985 ms; M_{Hole_fill}=2088 ms, SD_{Hole_fill} =1179 ms) compared to the SLD (M_{Hole_add}=1873 ms, SD_{Hole_add}=655 ms; M_{Hole_fill}= 1757 ms, SD_{Hole_fill} = 469 ms), t_{Hole_add} (19) = -1.556, $p> .005$ and t_{Hole_fill} (19) =-1.371, $p>.005$.

A t-test analysis was done to verify whether the head movements had an effect on response times on images presented on the MLD, the effect was not significant, $t(11)$ =.446, $p>.05$.

6.4.2.2. Accuracy

The main effect of the display type was significant, $F(1,22)=16.57$, $p<.01$. The main effect of the manipulation type was also significant, $F(5,110)=57.71$, $p<.01$. A closer inspection to Figure 6-3 might suggest that the luminance differences were the most difficult to detect (M=.55) together with the hue differences (M=.75) while the other differences provoked detection rates greater that 80%. However further analyses is required.

[10] Post hoc analyses were conducted using the Bonferroni adjustment based on ten comparisons, therefore α=.005

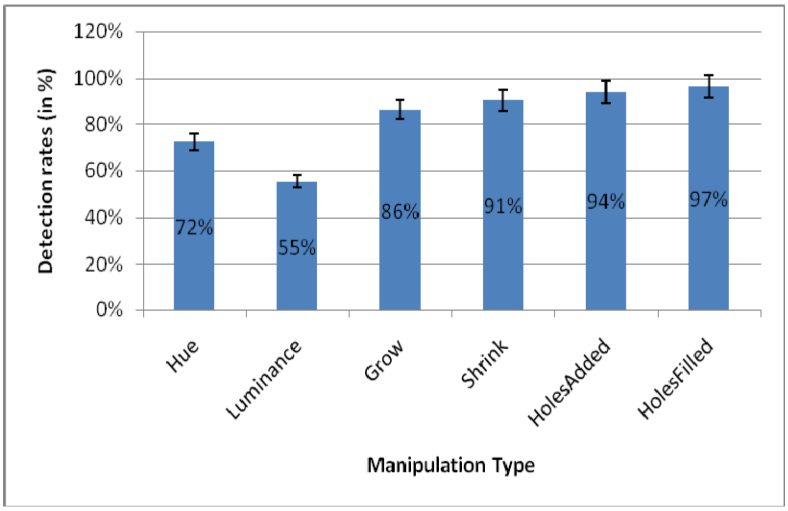


Figure 6-3. Detection rates per manipulation type (Vertical bars depict 95% confidence levels about the mean in each condition).

The interaction effect violated the sphericity assumption but the results were significant with and without the correction, $F(5,110)=25.37$, $p<.01$. Post hoc analyses indicated that participants were significantly more accurate when hue differences where presented on the SLD ($M = .90$, $SD = .15$) compared to the MLD ($M = .63$, $SD = .15$), $t(19) = 6.242$, $p<.005$. However, there was no significant difference in accuracy on the detection of "holes" presented on the SLD ($M_{Hole\_add}=.98$, $SD_{Hole\_add}=.05$; $M_{Hole\_fill}=.99$, $SD_{Hole\_fill}= .04$), or the MLD ($M_{Hole\_add}=.94$, $SD_{Hole\_add}=.09$; $M_{Hole\_fill}=.96$, $SD_{Hole\_fill} =.08$), $t_{Hole\_add}(19)=1.753$, $p> .005$ and $t_{Hole\_fill}(19) =2.179$, $p>.005$.

It was also assumed that luminance differences in the MLD ($M= .09$, $SD=.15$) would not be noticeable provoking a high rate of non-responses compared to the SLD ($M=.13$, $SD =.13$). Paired t- test resulted in non-significant differences, $t(19)=.767$, $p>.005$. Nevertheless, as mentioned above, the detection of luminance seemed to be negatively affected and it was found that participants were more accurate with the side by side presentation ($M=.79$, $SD=.17$) compared to the detection of the same differences when the images where superimposed ($M=.34$, $SD=.19$), $t(19) =7.285$, $p<.005$, suggesting that the MLD impaired perception of luminance differences causing a higher degree of error.

A t-test analysis was conducted to verify whether the head movements had an effect on accuracy on images presented on the MLD, the effect was not significant, $t(11) = .527$, $p > .05$.

6.4.3. **Conclusions**

The first hypothesis indicated that hue differences in the MLD would provoke slower and less accurate responses than in the SLD. In fact, hue differences were detected more accurately when images were presented side by side. However, the difference between response times was non-significant.

The second hypothesis stated that luminance differences in the MLD would not be noticeable provoking a high rate of no-responses compared to the SLD. The rate of non-responses between the SLD and the MLD yielded non-significant results. An analysis of the detection rates showed that the detection of luminance was significantly more accurate when the images where presented side by side compared to the detection of the same differences when the images where superimposed suggesting that the MLD impaired perception of luminance differences causing a higher degree of error.

The third hypothesis indicated that 'holes' differences in the MLD would provoke slower and less accurate responses than in the SLD. Contrary to our hypothesis, the results showed otherwise suggesting that the MLD did not impair the detection of an obvious difference.

Finally, it was hypothesised that changes in size (grow and shrink) would provoke faster response times in the MLD. Indeed, differences in size were detected almost four seconds faster when superimposed compared to their coplanar counterpart.

Thus, differences in size are detected faster and more accurately when superimposed than when presented side by side but more subtle differences such as changes in hue or luminance are better detected when presented on a single layer display.

6.5. Experiment 4: Comparing complex stimuli

Participants had a total of 6 practice trials and 126 experimental trials whereas the session lasted an average of 30 minutes. For a short video of this experiment see Appendix III – Video 4.

Six hypotheses were tested:

1. Translations in the MLD would be detected faster and more accurately than in the coplanar presentation.
2. The MLD would be no better than the coplanar (or side-by-side) presentation for the detection of addition of objects.
3. The MLD would provoke faster response times for the detection of deletion of objects than the SLD.
4. Deletion in the MLD would be detected faster and more accurately than addition in the MLD.
5. Participants will be slower and less accurate detecting differences on photographs on the MLD than on the SLD.
6. Simple images with a white background (shapes) would provoke faster and more accurate responses in the MLD than images with coloured backgrounds (photographs).

6.5.1.1. Stimuli

126 pairs of images were randomly ordered using a latin-square design. There were four types of images: shapes, cartoons, Excel graphs and photographs. Figure 6-4 shows a sample of the images used for the single-layer condition. The red rectangles highlight the difference, the arrows indicate translation. Arrows and rectangles were not presented during the experiment.

All stimuli presented three main types of differences:

- *Addition* – one object was inserted to the image,
- *Deletion* - one object was removed from the image, or
- *Translation* - one object was shifted, say for example from right to left, compared to the object in the original image.

The data for "deletion" of cartoons in the SLD condition were missing for the first eight participants. There were also some data points missing because the analysis was done based on correct responses only. Although the literature suggests several approaches for the treatment of missing data such as pairwise deletion, substitution with a measure of central tendency or imputation methods like the Maximum Likelihood and Multiple Imputation (Howell, 2009b; McKnight *et al.*, 2007), there is no universal rule of thumb for deciding whether or not to drop cases with missing values, or to impute values to replace missing values. When samples are large and the proportion of cases with missing data is small, it is common simply to drop these cases from the analysis (Gerber, 2005; Little *et al.*, 1987). In this case, the researcher decided to drop the missing cases.

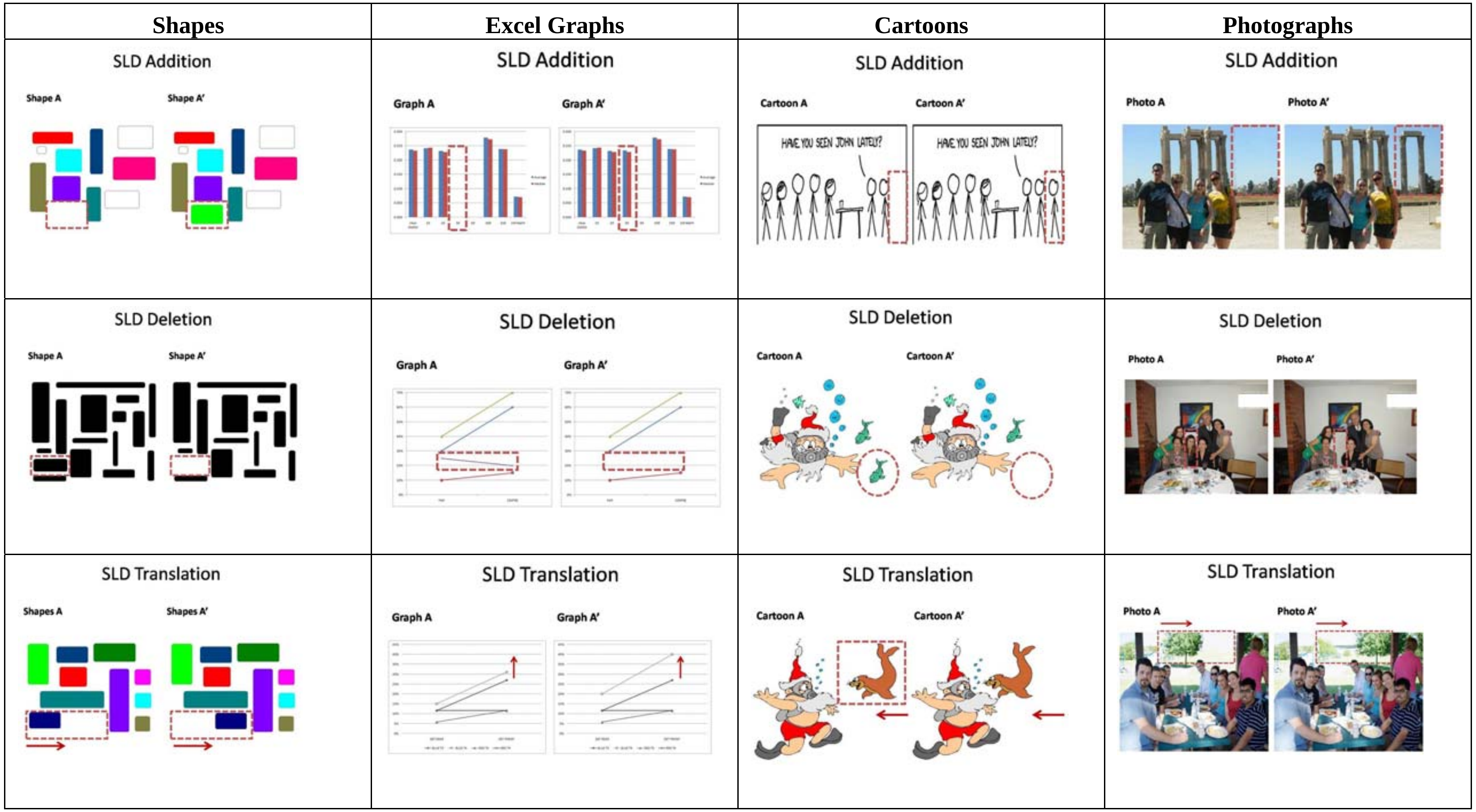


Figure 6-4: Experiment 4 - sample of images SLD condition.

6.5.2. **Results**

6.5.2.1. Response Times

Response times were subjected to a 2×3×4 repeated measures ANOVA having two levels of display type (SLD, MLD), four levels of stimuli type (shapes, cartoons, graphs, photos), and three levels of manipulations (addition, deletion, translation). For a full report of SPSS tables, see Appendix IX.

The main effect of the display type yielded an F ratio of $F(1,16)=6.389$, $p<0.05$. Post hoc[11] analyses indicated that participants' response times for the SLD presentation (M=4646 ms, SD = 742 ms) were significantly slower than the MLD (M=3429 ms, 1068 ms), $t(23) = 5.026$, $p<0.004$.

There was also a significant main effect of the manipulation type, $F(2,32)=10.302$, $p<0.01$. The main effect of the stimuli type yielded an F ratio of $F(3, 48)=49.556$, $p<.01$. The interaction effect between the display type and the manipulation yielded an F ratio of $F(2,32)=18.013$, $p<.01$. A closer inspection of Figure 6-5 suggests that the deletion of an object was detected faster than addition and translation differences.

However, post hoc analyses indicated that the difference between the detection of addition differences on the MLD (M=4284 ms, SD= 1281 ms) and the SLD (M= 5009 ms, SD =1186 ms) was not significant, $t(23)= 2.130$, $p>.004$. The difference in response times detecting the deletion (M= 3723 ms, SD= 1310 ms) and addition (M= 4284, SD= 1281 ms) of differences on the MLD was also non-significant, $t(23)= 2.255$, $p>.004$.

The detection of deletion was borderline significant and participants seemed to detect deletion faster on the MLD (M= 3723 ms, SD= 1310 ms) than on the SLD (M= 4606 ms, SD= 1011 ms), $t(23)= 3.125$, $t(23) = 3.125$, $p=.005$. Translation differences, on the other hand, were detected significantly faster

[11] Post hoc analyses were conducted using the Bonferroni adjustment based on thirteen comparisons, therefore $\alpha=.004$

on the MLD (M= 3598 ms, SD= 1321 ms) compared to the SLD (M= 6808 ms, SD= 2230 ms), t (23)= 6.125, p < .004.

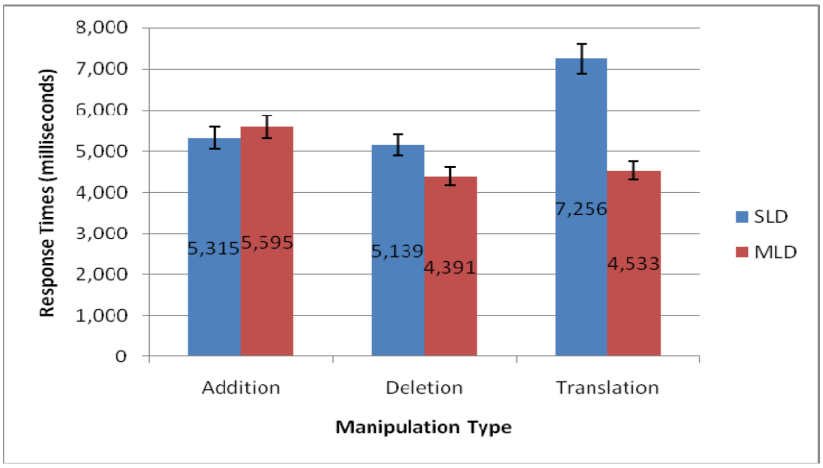


Figure 6-5. Average response times in ms based on manipulation and display type (Vertical bars depict 95% confidence levels about the mean in each condition).

There was a significant interaction between the display type and the stimuli type, F(3,48)=9.822, p<.01. By examining Figure 6-6 it seems that participants detected differences in shapes, cartoons and photographs faster when presented in the MLD compared to the side-by-side condition. Post hoc analyses indicated that differences on shapes presented on the MLD (M=2788 ms, SD= 992 ms) were detected significantly faster than the photographs' differences (M= 6694 ms, SD= 2381), t (23)= 9.138, p <.004. The difference in response times between the detection of differences on photographs presented on the MLD (M= 6694 ms, SD= 2381 ms) compared to the SLD presentation (M= 7268 ms, SD= 1863 ms) was non-significant, t(23)= -.969 , p > .004.

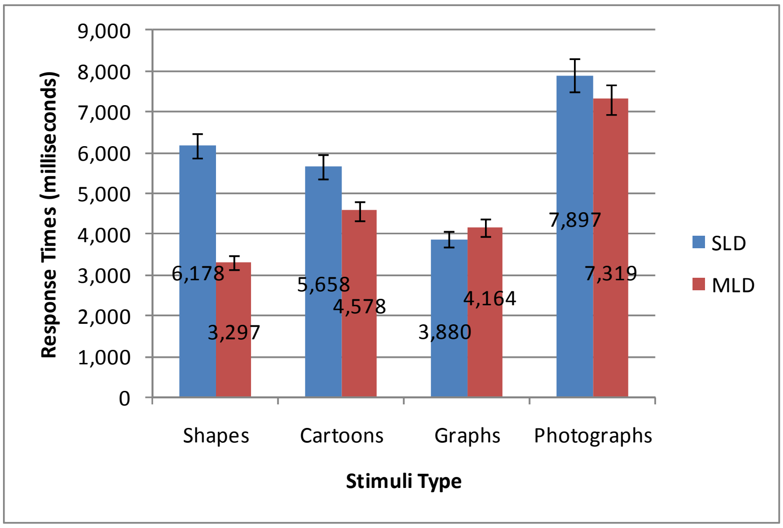


Figure 6-6. Response times by stimuli and display type.
(Vertical bars depict 95% confidence levels about the mean in each condition).

There was a significant interaction between the stimuli type and the manipulation, F(6,96)=4.355, p<.01. The interaction violated the assumption of sphericity but the results were significant with and without the correction. Hence the F ratio reported above is the one without the correction. Finally, the interaction between the display, the manipulation and the stimuli type also violated the sphericity assumption but F was significant with and without the correction, F (6, 96) = 3.498, p< .05.

6.5.2.2. Accuracy

The main effect of the display type was not significant, and neither was the main effect of the manipulation type. The main effect of the stimuli type yielded a significant F ratio of F(3, 69)= 64.85, p<.05.

The interaction effect between the display and the manipulation type yielded an F ratio of F(2,46)= 7.702, p<.01. Figure 6-7 suggests that the deletion of items was better supported when images were presented side by side while the detection of translation differences was more accurate when presented in the MLD. Post hoc analyses showed that participants were significantly more accurate when detecting translation differences on the MLD (M=.82, SD=.13) than on the SLD (M=.66, SD=.11), t (23) = -4.296, p <.004. They were also significantly faster when detecting addition differences but on the SLD (M= .84, SD =.11) compared to the MLD (M= .75, SD= .12), t (23) = 3.334, p< .004. there was no significant difference in accuracy between the detection of deletion (M= .79, SD= .11) and addition (M= .75, SD= .12)when presented on the MLD, t (23) = -1.472, p> .004.

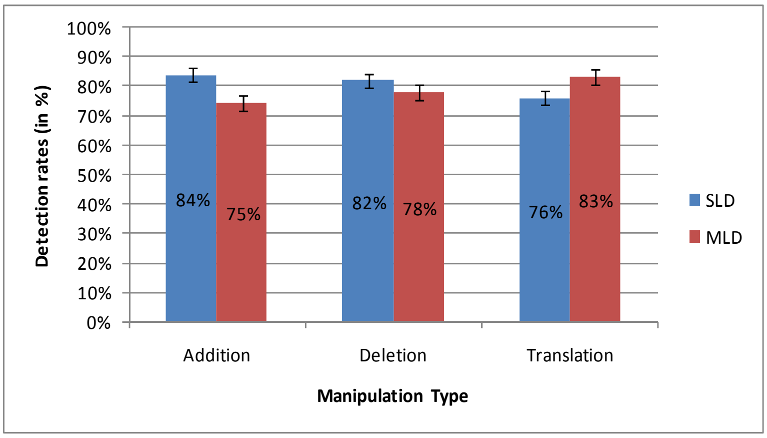


Figure 6-7. Detection rates by manipulation and display type (Vertical bars depict 95% confidence levels about the mean in each condition).

The interaction between the display and stimuli type violated the sphericity assumption, but it yielded a significant F ratio with and without the correction, F (3, 69)=17.561, p<.01. Post Hoc analyses showed that differences in shapes were detected very accurately on the MLD (M=.84, SD=.09) but the detection of differences in photographs was impaired by the coaxial presentation of the MLD (M=.52, SD= .15), t (23) = -9.805, p<.004 (Figure 6-8).

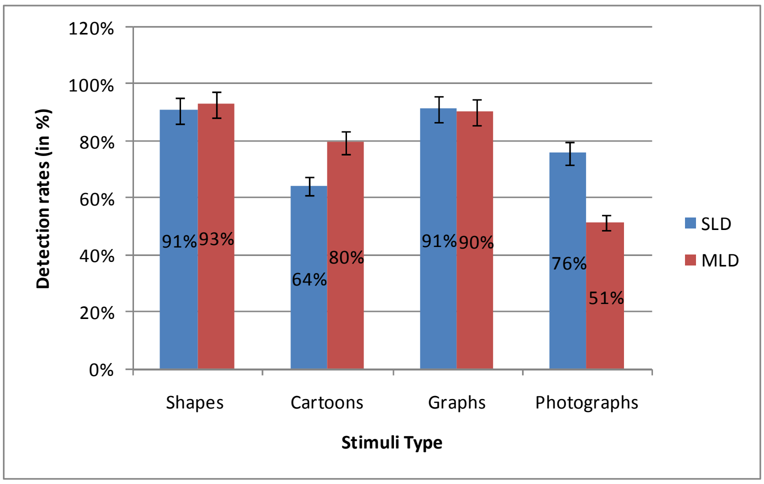


Figure 6-8. Detection rates by stimuli and display type.
(Vertical bars depict 95% confidence levels about the mean in each condition).

The interaction between the manipulation and the stimuli type was significant, $F(6,138)=9.359$, $p<.01$. Finally, the interaction effect between the display, the manipulation and the stimuli type violated the sphericity assumption but yielded a significant F ratio with and without the corrections, $F(6,138)= 5.802, p<.01$

6.5.3. **Conclusions and Limitations**

The first hypothesis predicted that that the detection of translation differences in the MLD would be faster and more accurate than the SLD. Results suggested that participants were significantly faster and more accurate detecting translation differences on the MLD compared to the coplanar presentation.

The next hypothesis suggested that the MLD would be no better than the SLD condition for addition of objects but there was no significant difference. However, response times for detection of deletion on the MLD was significant compared to deletion of objects on the SLD.

The detection of differences in complex stimuli such as photographs were detected more accurately when presented side by side compared to the coaxial presentation but the response times were not significantly different.

Participants detected the differences of simple images with a white background, such as shapes, faster than complex stimuli such as photographs when presented on the MLD. Accuracy was also significant but the differences in complex stimuli were detected more accurately presented side by side.

This study exposed one limitation because the detection of translation differences could have been due to an effect of colour instead of depth. If this colour effect is mimicked on a single layer screen by using alpha blending and monocular cues, the possibility that the detection of translation differences was due to colour perception could be eliminated.

6.6. Discussion and Conclusions

The purpose of these experiments was to explore the utility of the MLD for detecting differences between two states. Two experiments were designed to evaluate participants' ability to detect differences of simple and complex stimuli such as shapes or photographs.

With simple stimuli, participants were faster and more accurate detecting changes in size and shape when the images were presented on the MLD compared to the side by side presentation.

It was hypothesised that participants would be faster and more accurate to detect translation and deletion of objects on the MLD than the SLD condition. Yet, only the detection of translation differences showed to be faster and more accurate on the MLD than on the SLD.

Overall, the MLD provoked the fastest response times but with significantly higher error rates than the SLD condition. Additionally, and contrary to our hypotheses, deletion of an object in the MLD was not detected faster than the SLD condition.

Having hypothesised that the MLD would provoke poor performance in the photographs condition due to the clutter and complexity of the visual scene compare to the SLD condition, it was surprising that the results indicated that it took equal times to detect the difference in photographs in each condition, but detection was possible far more often in the SLD than in the MLD condition. A similar situation occurred with the hues and luminance differences in which the MLD did not affect response times, but rather the capacity to perform the task.

It is important to note that although the concept of a *difference* in a multi-layered display differs from the modifications that have been used in normal (single-layered) displays, the cognitive limitations manifested in similar ways. As mentioned in section 3.2 and 6.1, it is assumed that some mental representation of the visual environment is maintained across eye movements in a trans-saccadic memory but the information derived from a single fixation is usually insufficient to base comprehension (Irwin *et al.*, 1998; Parkin, 2000). The concept of addition and deletion requires a representation of each layer and direct comparison to detect the existence or non-existence of an item in one of the layers. It means that the comparison between layers is also affected by saccadic suppression in the same way it affects detection of differences on single layer displays when images are presented side by side. Participants have to go back and forth from one layer to the other to detect the difference on the MLD, while on a single layer display; their eyes go side by side. From previous research we know that the comparison process on single layer displays takes about 20ms/item if the items are already in short term memory (Rensink, 2002). Further research is required to measure the time that the comparison process takes per item when using a multi-layer display. Nevertheless, it seems safe to say that the only instance where saccadic suppression did not affect participants was when using the MLD to detect translation differences. In those cases, the comparison process with a multi-layered display was immediate and direct. Hence, results suggest that the MLD has limited application in the field of change detection, since only the detection of translation differences and changes in size resulted in important significant differences but, within those limits, the MLD has the potential to be a powerful tool.

Chapter 7

'CHANGE BLINDNESS' IN THE BRITISH TRANSPORT POLICE

The literature suggests that failure to detect changes is much more probable if the change occurs coincident with a distraction or interruption (Durlach, 2004b; Smallman *et al.*, 2003) such as a popup menu that blocks the change, or when working with multiple screens to the point that multitasking can slow change detection (Di Vita *et al.*, 2004; Durlach, 2004a, 2004b; Richard *et al.*, 2002; Smallman *et al.*, 2003; St. John *et al.*, 2007). Previous studies that have investigated change blindness in operational environments have used simulations of graphical-tactical displays to monitor navy, army or air traffic (Di Vita *et al.*, 2004; Durlach, 2004a, 2004b; Durlach *et al.*, 2008; Muthard *et al.*, 2002; Richard *et al.*, 2002; Smallman *et al.*, 2003; St. John *et al.*, 2007; Wickens *et al.*, 2003) but they have not attempted to evaluate change blindness in situ. This suggests that a study to undertake is one in the field to investigate the change blindness phenomenon in a control room where change blindness may occur.

Access was obtained to the British Transport Police (BTP) Force Control Room in London (FCRL), United Kingdom. The field study explored the change blindness phenomenon in the BTP control room to determine whether operators miss changes when multitasking or after an interruption.

7.1. British Transport Police (BTP)

In the United Kingdom there are several independent police forces. Some of these forces operate within specific geographical territories, such as the various counties. Others have responsibility for particular areas of activity, such as the BTP, the specialist force for Britain's railways. They monitor the journeys of over six million passengers and 400,000 tons of freight daily along 10,000 miles of track.

BTP operates two control rooms and one call-handling centre. The *First Contact Centre* is responsible for handling all routine telephone traffic. The *Force Control Room – Birmingham (FCRB)* is responsible for the East Midlands, West Midlands, Wales, the North West of England, the North East of England, the South West of England and Scotland. The *Force Control Room – London (FCRL)* is responsible for the Greater London area (including the London Underground and Mainline), London North and London South areas which are usually known as the Home Counties. Figure 7-1 shows a map of England where the coloured areas represent the BTP FCRL operational areas: London North in light green, London South in dark green and the London Underground in dark blue.

Figure 7-1. FCRL operational areas

The FCRL's goal is to provide seamless communications to the three policing areas of the mass transit system; London North, London South, and the London Underground, covering more than 10,000 miles of track.

The FCRL's core tasks are to direct, monitor and support police resources; provide first-instance supervision of the policing incidents until a supervisor arrives at the scene; and to record police actions. There are two requirements for command and control: to establish the resources and capability required and the co-ordination of their deployment in an emergency.

The FCRL physical layout is divided into four areas: call takers, radio dispatchers, CAD (Computer-Aided Dispatch) operators and the supervisors' top desk. The monitoring and control of units is carried out in the dispatch area. CAD operators monitor messages that come from the Metropolitan Police and the London Ambulance Service into a different system from the one the BTP uses. The FCRL is divided into three sectors mirroring the operational sectors: London North, London South and London Underground (Figure 7-2).

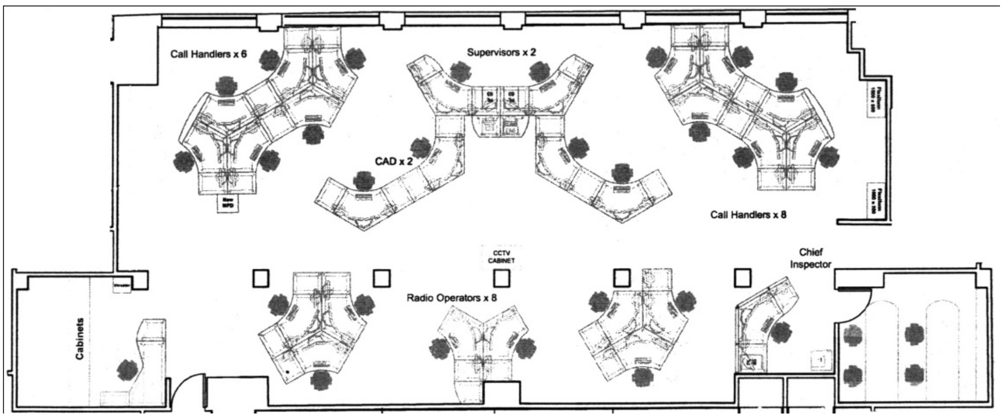


Figure 7-2. Floor Plan of the FCRL

After interviewing call takers and CAD operators, the researcher decided to focus the analysis on radio dispatchers because they coordinate resources, manage the incidents and use between three to five monitors.

7.1.1. **Radio dispatchers' job**

When a call is received, a call taker obtains the details, determines the incident classification and its grade or priority. 'Immediate' incidents (abbreviated in the system as IMM) are referred as G1 incidents and units will use their sirens. 'Prompt' incidents (PM) are referred as G2 and units are required as soon as possible but they cannot use their sirens. "Routine" incidents (RO) require police assistance but officers can take longer to respond.

When it has been determined that an incident needs to be resourced, the call taker sends the information to a radio dispatcher by an action message. Once the action message is received, the dispatcher reviews the description of the incident, reviews available police units, assigns an appropriate number of units, and monitors ongoing incidents and outstanding jobs. The dispatcher may decide to change the priority classification. In addition, the dispatcher receives reports directly from officers in the field regarding new and ongoing incidents. New incidents reported by officers are graded as "Police generated" (PG). Information received from officers is logged into the command and control system known as NSPIS (National Strategy for Police Information Systems). The dispatcher is also responsible for checking on officers periodically after they have been assigned to an incident.

To assign available units, radio dispatchers usually "put a call out" which means that they call all units over the radio indicating grade and incident location. Units that are available and close to the incident location reply with an estimated time of arrival (eta). However, there are cases in which no units reply and radio dispatchers have to check their resource list (in NSPIS) (Figure 7-3). The resource list enumerates all units and radio dispatchers choose from those that have an activity code 01 which means that they are on duty and available.

Call Sign	Pw Ref	No	Location	Proximity	Activity code	Comment	Beat id	Resource type	Deployed	End time
tHM543	7890	1	Victoria		01	CAT		PCSO		23:00
tBK890	4356	1	King's Cross		02	CORE TEAM		PCSO	Deployed	15:00
LE51	0039	1	West Ham		51	DUTY OFFICER		Duty Sergeant	Deployed	09:00
B76		1			01			Vehicle		14:00
AX54		1			09			Drug dogs		23:00
tHA43	1209	1			02			Police Officer	Deployed	15:00

Figure 7-3. Schematic representation of the Resource List in NSPIS

7.2. Methodology

7.2.1. **Data Collection**

The field study comprised a total of 42 hours of observations, 12 interviews, and a review of incident logs and performance indicators (from April 2007 to April 2008). The first 15 hours of observations were spent with radio dispatchers to determine the frequency with which they shifted their attention between the different computer displays. The rest of observations were intertwined with the interviews and the main objective was to observe radio dispatchers as they went about their duties.

For the first 15 hours of observations, the researcher stood behind the radio dispatcher in order to have a clear view of monitor shifts. For the rest of the observations, the researcher sat next to the dispatcher with a set of headphones so actions and voice communications were easily monitored. Initially, the BTP Chief Inspector did not authorize any electronic equipment. He authorized a voice recorder for the last five interviews. Detailed notes were taken for all interviews and observations.

The interviewing method evolved over time. Initially, and very naively, the researcher enquired directly for a list of changes that operators usually miss. Operators reported that they did not miss any visual changes. This may have been because they do not actually miss changes, or because of the metacognitive phenomenon known as "Change Blindness Blindess" (Levin, 2002). This phenomenon explains why some observers are convinced that they would detect changes that others have missed during laboratory experiments.

The researcher then decided to interview operators about the use of the mapping tool, since previous research had found change blindness in graphical tactical displays (Di Vita *et al.*, 2004; Durlach, 2004b). After a few hours of observations, it was clear that operators did not use the mapping tool even though they were trained in its use. They reported that the map was too cluttered and was not efficient (See section 7.3.2.3 for further analysis of this point).

The researcher finally opted for a semi-structured interviewing technique known as the Critical Decision Method (CDM). The Critical Decision Method (CDM) was used to explore change blindness when operators were multitasking and after interruptions. The CDM is a retrospective cognitive analysis method that employs a set of cognitive probes to critical incidents that require expert judgement or decision making (Klein *et al.*, 1989; Wong, 2006). The CDM is generally used for eliciting expert knowledge, decision strategies and cues attended to in naturalistic decision making environments (Militello *et al.*, 2009). It has been applied in a number of domains including fire service, air traffic control, military and paramedics, naval warfare among others (Stanton *et al.*, 2005). The CDM has also been highly successful in eliciting perceptual cues and details of judgment compared to other traditional reporting methods (Klein *et al.*, 1989; Wong, 2006).

The CDM was used because it addresses the issue that demanding situations such as critical incidents do not occur on a regular basis and there was no guarantee that a suitable incident would occur while the researcher was in the control room. The retrospective nature of the CDM overcomes this problem by relying on experienced operators to recall a relevant incident. By focusing on non-routine or difficult incidents, the CDM produces a rich source of data about the performance of highly skilled personnel (Klein *et al.*, 1989) and uncovers elements of expertise that might not be found in routine incidents (Hutchins *et al.*, 2004). In addition, it is during critical incidents that they might encounter difficulties and challenging situations were technology could provide support for managing such incidents.

The CDM's flexibility allows the analyst to develop novel probes if the original probes are not adequate for the analysis (Stanton *et al.*, 2005), and in this study the probes evolved throughout the qualitative phase. For the final analysis, only the data from the two final sets of interviews were used. The first set used the original probes from Klein (1989) and determined whether operators missed changes when multitasking, especially during critical incidents. Based on the results from this set of interviews, the researcher decided to modify the probes and investigate interruptions which, according to the literature, provided an opportunity for change blindness to occur. The second set of probes focused on cue identification, situation assessment, and interruption recovery in order to determine whether operators were blind to changes after an interruption (Table 7-1).

Table 7-1. CDM probes to evaluate change blindness when multitasking and after an interruption

Multitasking during a critical incident: original probes (Klein, 1989)		After an Interruption: probes used for interruption recovery analysis	
Probe Type	**Probe Content**	**Probe Type**	**Probe Content**
Cues	What were you seeing or hearing?	Cue Identification for Interruption Recovery	What features were you looking for when you resume your task? How did you know that you needed that information to resume your task? Could you prioritise which piece of information is most important when resuming the task?
Information	What information did you use in making this decision or judgment? What are the action lists that you use? How and where did you get this information, and from whom? What did you do with the information?	Information Integration	What was the most important piece of information that you used to resume your task? What are the action lists that you use? How and where did you get this information, and from whom? What did you do with the information?
Analogues	Were you reminded of any previous experience? What about that previous experience that seemed relevant for this case?	N/A	N/A
Goals and priorities	What were your specific goals and objectives at the time? What was the most important to accomplish at this point in the incident?	N/A	N/A
Options	What other courses of action were considered or were available to you? How was this option chosen, and others rejected? Was there a rule that you were following in choosing this option?	Heuristics	Are you using any shortcuts or rules of thumb?
Experience	If a person of less experience than you were to face the same situation, what mistakes would he/she be likely to make?	N/A	N/A
Assessment	Suppose you were asked to describe the situation to someone else at this point. How would you summarize the situation?	Situation Assessment for Interruption Recovery	What information did you have available after the interruption?
Mental models	Did you imagine the possible consequences of this action? Did you create some sort of picture in your head? Did you imagine the events and how they would unfold?	N/A	N/A
Decision making	What let you know that this was the right thing to do at this point in the incident? How much time pressure was involved in making this decision? How long did it take to actually make this decision?	Decision blocking	Was there any stage in which you find difficult to process and integrate the information available after an interruption? Describe precisely the nature of the situation

7.2.2. **Data Analysis**

The interviews were transcribed and uploaded in *Hyper Research* (See Appendix X for full Transcripts). This software was used to assist in the coding and analysis process. The data analysis process involved the following steps:

1. Creation of decision charts
2. Summary of each incident
3. Creation of decision analysis tables
4. Identification of items of interest (Emergent Themes Analysis Approach)
5. Index and restructure each theme
6. Synthesis

The Emergent Themes Analysis approach explores the data to identify ideas and their relationships. It is based on Grounded Theory but tailored to take advantage of the exploratory and efficient data collection features of the CDM (Wong *et al.*, 2002). Wong and Blandford (2002) described it as a 'concept distillation' process that systematically identifies broad themes, and then iteratively refines the themes into more specific sub-themes that emerge from the data. Once the specific themes have emerged, the data are then categorized and summarized (Militello *et al.*, 2009; Wong *et al.*, 2002).

7.2.2.1. Decision charts, incident summaries and decision analysis tables

A decision chart is a visual representation of the process and the choices made during an incident showing the sequence in which the events occurred (Wong, 2004). An example of a decision chart is shown in Figure 7-4. The incident summary supplements the chart providing a description of the incident and uses the decision chart to organize relevant details from the transcript into an account of the stages of the incident.

The decision analysis framework shows how an operator is presented with situational cues and information, shows how this information is processed and states the immediate reasons for the action (Wong, 2004). An example of a decision table is shown in Table 7-2 (Appendices XI, XII and XIII show six incidents and their respective Incident Summaries, Decision Charts and Incident Timelines).

7.2.2.2. Identification of items of interest

This stage comprised the collation of the data obtained from observations and interviews. From the observations, the researcher obtained an estimate of the operators' shifts of attention from one monitor to the other and the number of interruptions that an operator faces. From the interviews' data, the researcher determined the visual changes that radio dispatchers have to detect, the cues – visual or auditory- that they attend to, and the strategies developed to identify those visual changes.

Liverpool Street – London Bombings (07/07/2005)

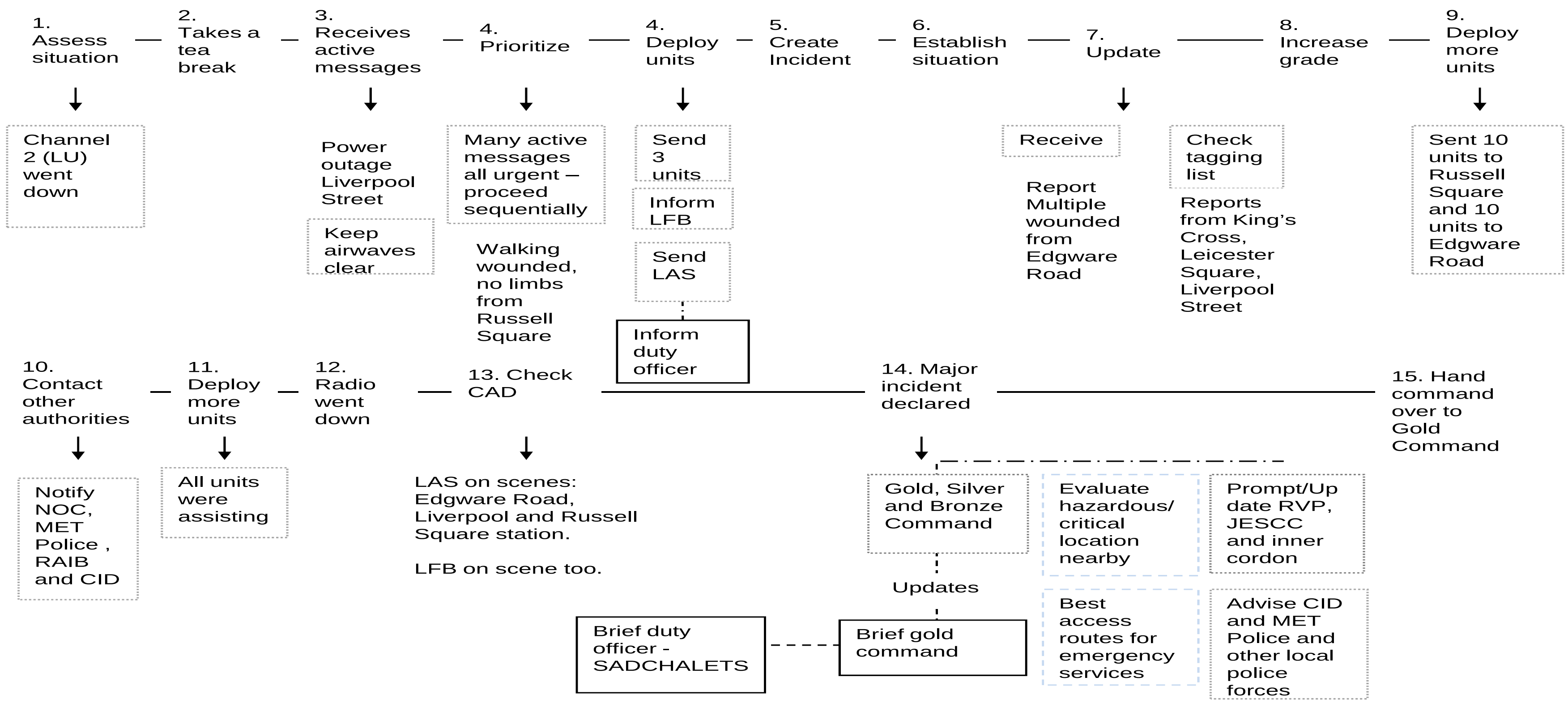


Figure 7-4. Sample decision chart – Liverpool Station Incident (interruption recovery)

Table 7-2. Sample Decision Analysis Table –Liverpool Station Incident (interruption recovery)

CUES	SITUATION ASSESSMENTS	ACTIONS	WHY?	WHAT FOR?
Action message (red message in the action queue)	Power outage Liverpool Street	Deploy 3 units	An officer on scene is needed	To report what is really happening on scene
Radio update - Call from officer on scene (ICCS)	Smoke or dust report Something suspicious	Call duty sergeant, LAS, MET Police		
Overheard another radio dispatcher	Bomb Damage Edgware Road	Deploy more units to both Liverpool and Edgware Station		To provide assistance
Other dispatch queues (when monitoring G8)			To keep situation awareness	To know what the situation is when finishing G8 hand over
Tagging list	Multiple Injured, limbs missing, people covered in dust	Inform NOC	Major disruptions in Central London	To manage the disruptions, technically and with the Press
Receives radio call	Bus exploded at Tavistock Square	Try to deploy more units, but none available.		To provide assistance

7.3. Observations

7.3.1. **Visual changes radio dispatchers have to detect**

Determining the qualities of a "change" for the FCRL was crucial for this study. Previous studies on change blindness in operational environments have defined changes as those relevant to the task (Di Vita *et al.*, 2004; Durlach, 2004a; Durlach *et al.*, 2008; Muthard *et al.*, 2002; Wickens *et al.*, 2003). For instance, when monitoring air traffic, the participant had to detect if an aircraft changes speed or course. For the purpose of this study, *change* was defined as the variation of an event that could be presented visually and was relevant to the radio dispatchers' job. However, a general definition of operational change was extremely difficult since the relevance of a change is context and task-dependent.

For the radio dispatchers, important events which variation needed to be notified were incoming calls, new incidents, incidents' updates, incidents' location, if an incident needs to be resourced, units' availability and their location.

According to Di Vita *et al* (2004), change blindness is likely to occur after an interruption unless the interface specifically draws attention to the changed information when the operator resumes viewing the previously unattended display. The researcher observed that appropriate visual and auditory cues were used to signal changes about incidents and units. The next sections describe the visual cues used by the two main systems used by the operators: the ICCS (Integrated Communication Control System) and the command and control system known as NSPIS (National Strategy for Police Information Systems).

<u>7.3.1.1. Visual cues on the ICCS</u>

Since operators are heavily loaded with voice communications, the radio and telephony communication functions are presented visually on a single operator touch screen that integrates the control of the communication subsystems. This system is known as ICCS. Its interface displays a list of the incoming calls on the right-hand side of the display. It also displays the previous calls on the left-hand side which is called "the ladder". Figure 7-5 shows a schematic representation of the ICCS interface. The dotted lines indicate the interface areas that change colour or flash to get the radio dispatchers' attention.

The ladder indicates the officer's call sign or the affiliation name such as LU manager or DLR controller. When a call comes in, the officer's call sign appears at right-hand side list and flashes until the call is answered. The channel used during the call also gets highlighted in bright green and shows the affiliation name or the officer's call sign and radio sign. The call sign indicates the geographic area and the specialization of the officer while the radio sign is a unique number that has to be reported to confirm the identity of the officer in case of an emergency.

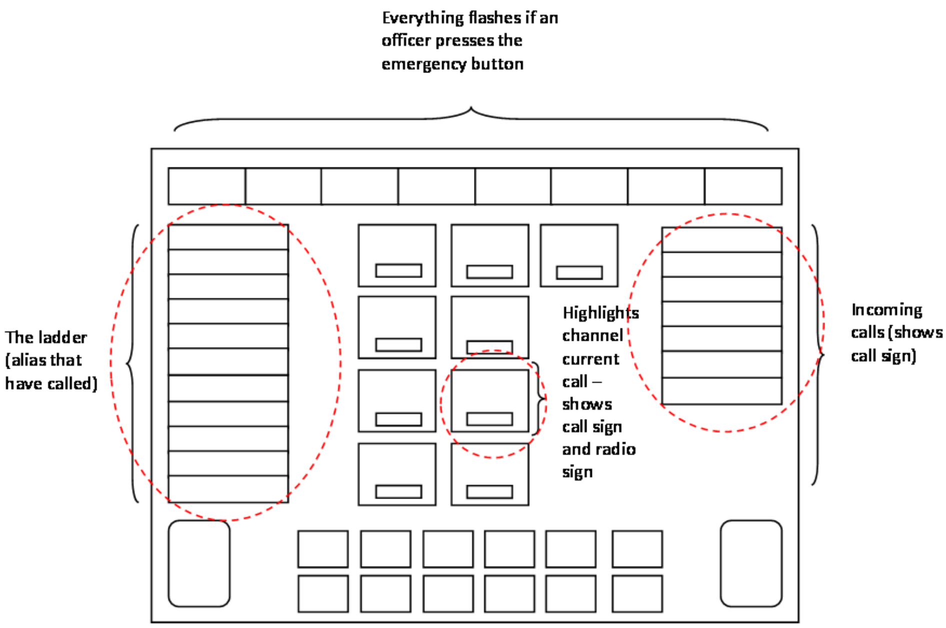


Figure 7-5. Schematic representation of the ICCS interface.

7.3.1.2. Visual cues on the NSPIS

The NSPIS command and control system is a text-based application that supports operators in creating, monitoring and recording incidents. This system provides the BTP with the platform to determine the resources required and coordinate their deployment. It provides radio dispatchers with a dispatch action queue that works like an email inbox (Figure 7-6).

New incidents that need to be resourced come at the top of the list coloured in red. Once acknowledged, the message turns black. Incident updates turn the message blue. If an incident has been acknowledged but not resourced, then the message turns green.

Incident No.	Grade	Description	Location	Units
345	IMM	TRAIN DERAILMENT	STEVENAGE	
432	PM	DISTURBANCE	VICTORIA	LS34, LX512
123	IMM	FATALITY	LEICESTER SQUARE	LA456,L56,LS756 ,L95
458	RO	TICKETING	KING'S CROSS PLATFORM 18	BD345,BE456
234	PG	DRUNKNESS	BOURNEMOUTH STATION	LE35
543	IMM	POLAC	GAINBOUROUGH ROAD	
765	DE	DRUG COLLECTION	SOUTHHAMPTON	H67, H57
834	PG	ROBBERY	IPSWICH	BG54
601	PG	TICKETING	OXFORD CIRCUS	BK95,BG153
249	PG		BANK	BK598,BG23
746	PG		LEIGHTON BUZZARD	B55,BG274

Figure 7-6. Schematic representation of the action queue in NSPIS

Thus, radio dispatchers are notified about a new incident with an incoming call (in the ICCS) or an action message in their action queue (in NSPIS). The incident location is found in the action queue or in the incident entry where the incident gets recorded. The units' availability is usually obtained when officers reply to a radio dispatcher's "call out", or if there is no reply it can be checked in the resource list code 01 (Figure 7-3).

7.3.2. **Observations: estimating possibilities for missing changes**

The literature suggests that shifts of attention from one monitor to the other can cause missing changes on unattended screens (Di Vita *et al.*, 2004). To form an upper bound on the number of changes an operator might miss, the researcher counted the frequency of monitor shifts. The researcher stood behind the radio dispatcher to count the monitor shifts. A monitor shift refers to the number of times an operator allocates his/her attention to a different monitor. Supposing an operator had three monitors, if the operator were to shift from monitor 1 to monitor 2, and then on to monitor 3 (M1 to M2, M2 to M3), two monitor shifts would have occurred.

Estimating these monitor shifts without an eye tracker posed limitations to the accuracy of the count. However, head movements and clicks were a good indicator because BTP radio dispatchers have specific functions per monitor. A radio dispatcher desk has three to five monitors (see section 7.3.1). The monitor on the left is the ICCS (telecommunications systems), the middle monitor is the NSPIS (command and control system), the monitor on the right is the one that has the map application but is usually used to check their email or the intranet. The CCTV[12] monitor is located on the far right and in order to check it the operator has to turn his/her body to see the screen. Radio dispatchers may also have a fifth monitor for accessing the CAD system. If that were the case, this monitor could be located between the NSPIS and the map monitor or on the far left. Because of this specific functionality, the researcher, by standing behind the radios dispatcher, could count the head or body movements and mouse clicks to estimate shifts of attention.

For instance, the radio dispatcher must read the call ID that appears in the ICCS monitor when a call comes in. Therefore the radio dispatcher will point at the touch-screen to answer the call and/or move his/her head to read the ID. While radio dispatchers are on the phone, they either write the details on

[12] Closed Circuit Television

paper, an easily noticeable action, or they click on the NSPIS monitor to create or update the incident; again the change in focus is easily noticeable. The interaction with the third monitor where they check their email and/or the intranet was not as noticeable as with the other monitors. In order to reduce confirmation bias and avoid situations where radio dispatchers' mouse was pointing on that screen but they might be looking at a different monitor, the researcher counted those instances where they clicked their inbox or swapped windows on that specific monitor. Interacting with the CAD or the CCTV required a body movement which was easily detected by the researcher. The researcher estimated that in the FCRL, on average, three changes of monitor per minute occurred during low/medium workload periods.

It was also observed that not only do radio dispatchers shift their attention from one monitor to the other; they also shift their attention from one incident to the other by opening new windows. Shifting their attention between monitors potentially creates an opportunity for radio dispatchers to miss changing information in the unattended displays. Additionally, opening incident logs in new windows occludes the action queue, creating an opportunity for change blindness.

The literature also suggests that interruptions in the workplace are analogous to the visual disruptions in change blindness experiments (Di Vita *et al.*, 2004; Durlach, 2004b). The researcher estimated the number of possible interruptions radio dispatchers deal with to assess how many times radio dispatchers could miss a change due to an interruption. Radio dispatchers usually have an average of five fifteen-minute breaks during an eight-hour shift which means that at least five times a day they interrupt their activities and have to be fully functional as soon as they return from their break. There are also intermittent interactions like briefing a supervisor, talking with a colleague about an incident or monitoring other action queues that involve interruptions. Each one of the conditions mentioned above is, according to the literature, sufficient to cause change blindness in an operational setting.

7.3.3. **Strategies to identify visual changes**

From the observations mentioned above, it is clear that radio dispatchers at the BTP have to deal with several interruptions and they have to multitask while managing multiple incidents that get constantly updated. One of these conditions, according to the literature, is enough for change blindness to occur. However, radio dispatchers in the BTP have developed a series of strategies to detect visual changes when multitasking and when recovering from interruptions in order to manage incidents successfully without an apparent effect from change blindness. These results suggest two crucial differences with previous studies. In the first place, previous studies that have analyzed the cognitive implications of the phenomenon have presented changes that are not relevant to the participant. The changes might be relevant to the gist of the scene but not for the participant. For instance, if a building disappears in a photograph or if a gorilla appears in the middle of a basketball game, it will affect the scene, but the existence of the stimuli is not relevant to the participant's life outside the laboratory. Second and most importantly is that even those studies that have used applied scenarios and have presented changes that are relevant to the task cannot give the participants the time to develop strategies to detect changing information.

Keeping in mind these differences, the following sections describe three main areas in which the BTP radio dispatchers have developed different approaches to deal with changing information: when multitasking, when they need to recover from an interruption and when handling geographic information.

7.3.3.1. Multitasking strategies

Radio dispatchers use three main strategies to detect changes when multitasking, they:

1. Log into other areas' action queues
2. Use their "control ears"
3. Prioritize calls

Although not strictly considered good practice, some radio dispatchers log into other areas' action queues to obtain a more complete picture of the situation. Radio dispatchers maintain their awareness by constantly screening the ongoing incidents even when they are doing a different task. For instance, a radio dispatcher who was working on July 7th 2005, the day of the London bombings, was monitoring the 31st G8 summit held in Scotland, but was logged into the action queue to monitor the incoming incidents.

> *R8: While I was in the radio with Glasgow [monitoring the G8 summit], I was still logged in the action queue monitoring the messages. I saw a call came in from Russell Square [station], this was different…(46-48).*

Radio dispatchers are able to listen to what the call takers or CAD operators are saying. They know that something is coming to them so when the action message arrives, they are ready to go.

> *R2: …That's why is important to have environmental awareness – when a call comes in; you know by listening to the call taker that something is coming to you (50-52).*

> *R9: It is just something that develops. I keep listening to both the radio and the environment. When you are first on the radio, you hear something on the radio but you won't hear anything else in the room, not even the person sitting next to you training you because you are concentrating so hard. But after some time, you can listen to the radio, type what they are saying, and listen to the call handler talking about the incident that you are dealing with, and look up. But you can have a conversation and hear the radio and hear to the other person (197-205).*

Radio dispatchers emphasised the importance of being aware of what is going on in the control room. They use their 'control ears' to listen for key information from others in the control room to corroborate information about incidents so they can update the incident log or make decisions about resourcing before an update is in the system. One dispatcher had managed an incident that involved a person under a train and severe delays in the lines coming to Euston Station in Central London. He reported that an unconfirmed incident started as a rumour from one of the passengers, which escalated into a serious accident:

> *R11: I said "hang on; we don't know if anything has happened". So I picked one unit, the BG [bravo golf] unit, but only on a G2 (18-19).*
>
> *Then, I heard, X speaking "Does anybody know where such and such bridge is?" Because I knew, I am from up there. My ears picked up at the sound of this bridge. So what have you got? I said. "I just had Network Rail Birmingham on telling me about an incident a person under a train". Ok, thank you, ok. So I went back to the radio. "Units now, now you can go" I get the units running again. Now it's confirmed (31-36).*

A dispatcher emphasised the importance of prioritization of calls regardless of the grading it has been given. They look at the grade and the incident type to make decisions about the type and number of resources needed in each incident.

> *R2: We have to prioritize. Any incident graded as a disturbance comes as a prompt. And more than often we need to send officers. But if an ambulance is required, it comes up as immediate. But if someone faints you don't need to send an officer (37-42).*

Another dispatcher in the day of the London Bombings also states the importance of prioritization, but indicates that on that day, he could only work in a sequential order because all messages coming through were real emergencies.

> *R10: I have all these active messages. They are underground; their active message is active to them. Everybody else was doing the same thing, for them they were active messages, they were urgent; they were trying to get it across (33-37).*
>
> *I wanted to prioritize the most important, so I said first unit stand by, second unit stand by, but they were all the same... So I did it in sequential order (38-40).*

7.3.3.2. Interruption Recovery Strategies

The radio dispatchers have two main strategies to recover from interruptions:

1. They tag only those ongoing incidents graded G1 (immediate) and G2 (prompt).

2. If interrupted, they go to the 'tagging list' and check for G1 incidents, their locations and units assigned.

Radio dispatchers reported that it was good practice to tag ongoing immediate or prompt incidents. They usually alternate between their action queue and the 'tagging' list (Figure 7-7). For instance, the radio dispatcher monitor the events during the London Bombings said:

> *R10: Then I checked the tagging list and back to my dispatch queue. Another active message "Walking wounded"... (32-33).*

The tagging list was initially developed to display outstanding jobs, but it is used for displaying ongoing incidents in each area. Tagging an incident is like using a flag in an email inbox but these "tagged" incidents are shown on a separate list. Radio dispatchers tag incidents which they are currently monitoring and could potentially escalate. This facilitates the search for specific incidents. For outstanding jobs, the system provides a reminder that rings at a predetermined time. Therefore, the 'tagging list' provides them with a concrete tool to rapidly visualise the current situation in a glance.

Incident No.	Grade	Description	Location	Units
345	IMM	TRAIN DERAILMENT	STEVENAGE	BK345,BG56,B43 , B67,AZ768
432	PM	DISTURBANCE	VICTORIA	LS34, LX512
123	IMM	FATALITY	LEICESTER SQUARE	LA456,L56,LS756 ,L95
458	RO	TICKETING	KING'S CROSS PLATFORM 18	BD345,BE456
234	PG	DRUNKNESS	BOURNEMOUTH STATION	
543	IMM	POLAC	GAINBOUROUGH ROAD	L35, LO987,LE45, L61...
765	DE	DRUG COLLECTION	SOUTHHAMPTON	H67, H57

Figure 7-7. Schematic representation of the "Tagging List" in NSPIS

> *R12: We tag it to the queue so that means that I can have a quick look of what is going on in the area, and just click that button there which is the Tag List Queue and just see instantly which messages are going through our queue. All the other messages are in the system at the moment, I'm not interested (168-172).*
>
> *R4: I can't do a search for every incident because it has been quite busy at the moment and I don't have a fantastic memory. So I have a look on the tags on the dispatcher queue just to be on the ground. Just to be tuned with dispatch. Then I can say "oh yeah, …" (64-67)*
>
> *R12: They [relief operators] will tell me only about ongoing incidents. Then, I flick through [the messages in the tag list] and read only G1 immediate calls (129-132).*

Radio dispatchers reported that the tagging list helps them to keep updated and did not state that they could have missed a change during an interruption.

> *R8: No, as I said, when I was assigned to be an AZ, I was logged in both queues… I couldn't miss a change. I knew what came from Russell Square… (59-60)*

> *R12: I wouldn't say you miss changes. When I came back to the radio, I said to the officer "Hang on a second". So when I came back, I was dealing with London south area, I can see control area which messages are in my area. I can also see [pointing to the tagging list] which messages are still open, London south A LSA. There might be a lot more calls that have been taken within this room, but these are the ones that we are dealing in the radio in this area (159-166).*

> *R10: Well…no, I don't think I missed anything. As soon as I came back, I went to the tagging list. At the same time, I was on the radio sending out the units. I knew what was happening, at least from Liverpool Station…(50 – 53)*

> *R10: As I said, went back to the main channel. Always look the tagging list, G1 and G2... (80-81).*

7.3.3.3. Spatial information management strategies

From the observations, the researcher noticed that radio dispatchers in the FCRL did not use the mapping tool available to them. This mapping tool displays the units on the map. It colour codes the units' availability. For those units that have been assigned to an incident, it displays the incident number next to the unit. In spite of the importance of geographic information to do their jobs, and the potential advantages that the mapping tool offers, radio dispatchers do not use it. Radio dispatchers preferred to use an online mapping tool called the *Gazetteer,* which works like Google maps, and only used when they need to give directions to officers. They reported that they do not use the mapping tool provided because it was too cluttered and it was impractical for their job.

> *R2: The map is used for giving directions only, we can zoom in and have a very detailed view of streets and small roads; but not to locate units; the map gets updated every so often, but we don't have time to sit there and watch it. It's impractical... (30-33).*

> *R7: The mapping system is not so good. For instance it doesn't differentiate the date. It is the national mapping system but it is not so user-friendly. We use the map for example to find addresses, but you can do that through Gazetteer which has live access to the station rail system (23 – 27).*

Previous research has only evaluated graphical tactical displays to measure and estimate the change blindness phenomenon. The fact that radio dispatchers did not use the map at all to monitor incidents meant that change blindness scenarios could not be identified from the use of this tactical display, and comparisons with previous studies were not possible. Instead of the mapping tool, radio dispatchers use the lists provided in the resource list by the NSPIS system to locate and assign units and find incidents' locations (See Figure 7-3). The CDM revealed that these tasks are inherently spatial but the way that operators perform them is currently non-spatial. Unit locations, availability, specialization and geographical specifications are listed in codes.

7.4. Discussion, Limitations and Conclusions

The funding for this project focused attention on evaluating the utility of the MLD for change detection tasks. The apparent need for a tool to enhance change detection was based on the assumption that change blindness is a problem in operational environments. Despite the several arguments about the occurrence of change blindness in operational environments and the consequences of missing operational changes, the results from this qualitative study suggest that change blindness did not occur at the BTP control room.

There are several possible explanations that require further research. The literature suggests certain conditions for change blindness to occur in operational environments. In the first place, the process of shifting our attention from one monitor to another creates an opportunity for changes to occur on unattended screens (Di Vita *et al.*, 2004). Second, interruptions in the workplace can also create an opportunity for changes to occur at the same time as the interruption. These interruptions could have a negative impact in tasks where operators monitor dynamic situations, from air traffic management to civil emergency response coordination, by disrupting users' situation awareness (Podczerwinski *et al.*, 2002; Smallman *et al.*, 2003; St.

John *et al.*, 2005; St. John *et al.*, 2007; Yeh *et al.*, 2000). Finally if operators have to multitask, they will have to divert their attention from one task to the other and therefore they will be more likely to miss a change on the unattended task. Multitasking and interruptions could affect the detection of changes in visual interfaces of complex visual displays where important changes in visually presented information could be missed if the changes occur coincident with a visual transient or distraction (Durlach, 2004b). Additionally, previous studies that have evaluated change blindness in operational environments have used graphical tactical displays.

This field study was conducted in the BTP's Force Control Room in London where operators:

- Have 3 or more monitors to work with.
- They were constantly interrupted.
- Have to manage several incidents at the same time.
- Have access to a graphical tactical display that has been referred as the "mapping tool".

With all conditions mentioned in the literature for change blindness to occur, it was safe to consider that change blindness will occur and might have consequences to the BTP operators' performance. From the observations at the BTP control room, it was estimated that three changes of monitor per minute occurred during low/medium workload periods. It was also estimated that operators interrupt their monitoring activities, at least, five times during a work shift when they go to a break. This minimum estimate does not include other non-scheduled interruptions which are sometimes part of their job like briefing a supervisor, talking with a colleague about an incident, or checking other action queues.

Given that there were opportunities for change blindness to occur, why were change blindness scenarios not found at the BTP? There are two main possible options: It could have been due to the limitations of the research method, especially if change blindness occurred, but did not impact performance. Alternately, a well designed interface, suited to the working environment evaluated, might have prevented change blindness from occurring.

It is quite possible that the CDM method is not sufficiently sensitive to identify situations in which operators were blind to changes. An alternative methodology to the CDM could have been employed. It is possible that situation awareness methods or error prediction methods could be more sensitive to change blindness. Instead of looking for incidents where change blindness might have occurred, the researcher could have begun by analysing incident reports where errors were already recorded, and then examining those errors where change blindness may have been a factor. A method such as SHERPA (Systematic Human Error Reduction) may have helped in the identification of those errors.

If we assume, however, that the radio dispatchers miss changes but were not aware of the fact, could be easily explained by a metacognitive phenomenon described by Levin (2002) known as "*Change Blindness Blindness*" were observers seem to believe that they would have no problem in detecting the changes that others have missed in laboratory experiments. However, if operators systematically miss visual changes, these errors would most certainly be addressed during their performance reviews and therefore they will be made aware of those errors. If they were aware of those errors, the CDM probes would have identified instances of the phenomenon. If on the other hand, they miss small visual changes that do not impact their performance, then change blindness does not matter because missing these changes is not operationally significant. So it could be possible that operators have learned to avoid missing important changes. A great example of this strategy is the use of the tagging list; after being interrupted, radio dispatchers check the tagging list for incidents graded as immediate only even though there might be updates on other incidents with a lesser grade.

The other possible option for not finding change blindness is a good interface design. The results from the interviews and observations made evident that both the communications and the command and control systems use appropriate visual and auditory cues to highlight relevant information. The command and control system provides a thorough logging that gives operators access on demand of the history of events. The 'tagging list' facilitates interruption recovery by allowing radio dispatchers to comprehend the current situation in a glance. Thus, the BTP system allows operators to go back not only to where they were before the interruption, but also allows them to check what happened while they were away.

It is also possible that change blindness could occur in an operational environment where a graphical-tactical display is used as the main workstation. Although geographical and tactical aspects of the BTP are clearly important, BTP radio dispatchers did not use the mapping tool reporting that it was too cluttered. In hindsight, graphical tactical displays seem susceptible to change blindness because changes are continuous and the nature of the state is dependent on multiple components. For instance, changes in air traffic control radar displays are dependent on a number of factors like speed, altitude and heading and the relationship between these factors. The pattern, therefore, has multiple components, if the traffic controller does not notice a small change in the pattern; it can escalate to become an incident very quickly. By comparison the radio dispatchers' displays are state-based and changes in these displays are discrete messages so a message is either update or not updated, a job is either new or acknowledged. Therefore, this last scenario allows for small changes that are not relevant to be missed since they are not integral part of any pattern and the rightness or wrongness of a state is far easily determined.

The fact that radio dispatchers do not use the mapping tool arises the possibility, although purely speculative, that the organization strategic planning enabled their operators to avoid a system where change blindness might have been detected. At this point, these are only speculations, and a comparison with other similar operational environments is necessary. Unfortunately obtaining clearance and access to such command and control centres is extremely difficult, hence, the lack of studies in the field.

These results support previous recommendations on the need for emergency services' command and control systems to have a thorough logging that give access on demand of the events' history (Malin *et al.*, 1991). For the BTP radio dispatchers, the tagging method facilitates the search of immediate ongoing events making it a very effective interruption recovery tool. Operators also have to go through a systematic training to become radio dispatchers. This training and the constant performance reviews allows them to achieve a higher level of expertise based on continued improvements in achievement. These improvements are not automatic consequences of more experience but a consequence of repeated *deliberate practice*[13]. This deliberate practice is the reason why they get the necessary experience to deal with demanding situations, to perform multiple tasks concurrently and to recover from interruptions.

In any case, these results call into doubt the hypothesis that change blindness is a problem in operational environments. Even though the theory suggests that operators that work with multiple screens, are frequently interrupted and perform multiple tasks simultaneously are vulnerable to change blindness, the field study did not identify any change blindness scenarios. This work was an exploratory first step. This field study is the first attempt to identify change blindness in the field as opposed to using simulations under laboratory conditions. This study could have been improved by focusing on the detection of unexpected events instead of the detection of visual changes. Changes, such as a new job in the dispatcher's action queue or an incoming call, are expected, but it still remains unknown whether the radio dispatchers ever see something unexpected on their screens and if they do, how important it is for their job performance.

The CDM was successful in determining strategies for successful job performance. The study identified that BTP radio dispatchers use several approaches to recover from interruptions, handle geographic information in a

[13] Deliberate practice is a regiment of effortful activities designed to optimize improvement (Ericsson *et al.*, 1993)

non-spatial way, and multitask. These results raised the possibility that the literature is based on an assumed, but not proven, causal link between change blindness and operational error. It also raises questions about how operators in operational settings handle changing information. It seems that they have learned to avoid missing important changes. When given the time to work with the system, they develop coping mechanisms. When considering radio dispatchers' interaction with the technology, these results suggest that a thorough logging, a "tagging list" functionality, appropriate use of colour coding and flashing, and setting one function per monitor seems to support change detection when monitoring dynamic situations. Finally, these results remind us that cognitive limitations found in the laboratory cannot always be found in the operational settings and that probably these cognitive phenomena are not a problem, just a consequence of how our attention works.

Chapter 8

CONCLUSIONS, REFLECTIONS AND FUTURE WORK

The central theme of this research has been the evaluation of the MLD as a tool to enhance the detection of visual changes. Given that empirical evidence has shown that we do not perceive everything that happens around us, even when paying attention to changing stimuli, many have assumed a link between change blindness and error in operational environments – especially when operators are multitasking or interrupted. Four laboratory experiments and one field study were conducted to determine whether the MLD is a useful tool in the kinds of change detection tasks that may be found in operational environments.

8.1. Is the depth of the MLD an efficient tool for change detection tasks?

The experimental results suggest that the depth cue of the MLD produces a positive effect in change detection tasks, albeit in limited conditions.

Experiment 1 showed that depth alone is sufficient as a visual cue to highlight changes. Participants' accuracy was superior to colour when changes occurred at a distance greater than five degrees of visual angle from the focus of attention. Results also indicated that the depth-transient may be detected as rapidly as a colour transient. Therefore, if depth is used as an alerting visual cue and observers are expecting the changes; depth could be used to highlight changes that occur at a distance greater than five degrees of visual angle from the focus of attention.

The results of Experiment 2 indicated that an unexpected event becomes more noticeable when it is highlighted with depth at a crossed disparity (similar to the front layer of the MLD), is close to the focus of attention, or presents is visually similar to the monitored objects. These results were consistent with previous studies were participants were more likely to detect an unexpected event if it was similar in colour, shape or luminance to the object they were monitoring (Most *et al.*, 2005; Most *et al.*, 2000; Most *et al.*, 2001; Simons *et al.*, 1999). Results from Experiment 2 also showed that the unexpected event becomes more noticeable when the rest of the stimuli are distributed in both layers and the unexpected event appears on the front screen.

Experiments 3 and 4 showed that, as a comparison tool, the depth of the MLD allowed fast and accurate detection of translation and size differences. A translation difference refers to an item in the image that has been displaced or moved. A size difference refers to an item in the image which dimensions have either increased or decreased compared to the original. The MLD did not help in the detection of missing objects, or the appearance of new ones, or of changes in luminance or hue. The detection of differences was also faster on the MLD if the images consisted of sparse, simple objects on a white background. In complex images, presenting the images side by side on a single layer provoked faster and more accurate detection rates than on the MLD.

Despite the apparently resounding success of the MLD for change detection, it is not yet recommended for use in an operational environment. In the first place, the experimental design of Experiment 1 did not properly control for motion. The changed digit popped up to the front layer producing a radial motion signal. No visual disruption was introduced to eliminate this signal. The higher accuracy obtained outside the parafoveal region could have been due to an effect of motion rather than depth. It is known that the peripheral vision is sensitive to motion due to its high density of rods. By replicating the study including a motion condition, say a condition were the digit vibrates, the issue of this possible confounding factor could be solved.

There is also the possibility that the lack of some conditions in Experiment 2 could have affected the power of the experiment. One could argue that positioning the stimuli in the rear layer for the control condition affected detection because the rear layer has the disadvantage of blurring the images slightly. Adding a condition where the stimuli were placed on the front screen would have partitioned out this possibility from the results. However both front and rear layer change the quality of the image each in a different way. While the image gets slightly blurred if presented on the rear layer, it becomes translucent if presented on the front layer. It is also important to note that during the control condition no other stimuli was placed on the front layer so factors such as occlusion or colour superimposition did not occur. More importantly, adding a new condition would have meant the sample size had to be increased at least by 60%. The current sample size is already rather small for any Inattentional Blindness study taking into account that each participant can only receive one inattention trial. Getting 60 people to help out with the experiment was difficult; getting 96+ definitely presents a challenge. With the current experimental design, there were some conditions left with only 5 data points. This sample affected the power of the experiment and although results showed strong levels of association for the main effects, the interaction between them provided inconclusive results.

Experiment 4 showed a significant effect of depth for the detection of translation differences. Nonetheless, there is a possibility that the detection of translation might be mimicked by superimposing colours on a single layer screen. Since monocular cues such as colour-blending and pictorial depth were not tested, the question remains unanswered and further research is necessary. Additionally, experiments 3 and 4 showed that the use of the MLD has to be limited to environments were the foreground and background images are not particularly cluttered.

In addition, the binocular disparity produced by the MLD is below the recommended minimum working value. The MLD comes with a fixed separation between screens that cannot be modified. The 17-inch MLD used in these experiments had a 14 mm separation between the screens. At a viewing distance of 630 mm, the binocular disparity of the MLD is only **8 arc seconds**. Previous studies on three-dimensional displays have suggested a minimum working value of 20 arc seconds. More recent research has found that for visual search tasks a minimum disparity of **6 arc minutes** is needed for participants to process depth in parallel and therefore preattentively. Because the binocular disparity of the MLD is below the minimum working level, most participants probably searched for changes in depth in a serial manner affecting their response times.

It is important to note that the characteristics of the display and the nature of human vision combine to define the limits of the overall system. Depth judgement performance cannot always be predicted from display geometry alone. Other system factors, including software drivers, electronic interfaces, and individual participant differences must also be considered when using a 3D display to make critical depth judgements. Based on the results from experiments 1 to 4, the depth of the MLD has an effect on detection but under limited conditions and unless the binocular disparity of the MLD is increased to at least the minimum working value, it is unlikely to be an efficient tool for change detection. Nevertheless, by understanding the characteristics of the MLD one could take advantage of both its limitations and capabilities. For instance, use the MLD as a binary tool placing important but simple stimuli on the front layer.

8.2. Is change blindness a problem in operational environments?

The field study explored the change blindness phenomenon in the BTP control room to determine whether operators miss changes when multitasking or after an interruption. The BTP control room met all the conditions described in the literature for change blindness to occur: Operators used multiple monitors, they were frequently interrupted, they tend to multitask and manage multiple incidents at the same time, and they have access to a mapping tool. Even with all the conditions met, the results from the field work call into question the applicability of change detection theories in operational environments. These results raised the possibility that the causal link suggested in the literature between change blindness and operational error may only have relatively minor consequences in command and control rooms similar to the BTP.

There are two possible explanations for not finding change blindness. Limitations of the research method, especially if change blindness occurred without impacting performance, may have led to a false negative. Alternately, a well designed interface, suited to the working environment evaluated, might have prevented change blindness from occurring.

It is possible that the critical decision method (CDM) was not sensitive enough to identify change blindness scenarios. If operators were not aware of missing any changes (consistent with the metacognitive phenomenon known as Change Blindness Blindness) then they would not have reported them in the CDM interviews. However, if they were persistently missing significant changes, these errors would have come up during their performance reviews, and operators would have been able to report these errors during the interviews. There is also the possibility that radio dispatchers missed changes, but those changes were small and therefore not detecting those changes did not have an impact in their operations. In this case, change blindness is not important. Nevertheless, alternative methodologies like situation awareness methods or error prediction methods such as SHERPA (Systematic Human Error Reduction) might have been more effective in detecting change blindness if it was occurring.

Most probably, operators have learned to avoid missing important changes. This is very different from missing other types of changes. Studies that have focused on the cognitive implications of change blindness present visual changes that might or might not be relevant to the gist of the scene and are not important for the participant's task. Remember, for instance, the gorilla experiment. The gorilla appeared among two teams that were playing basketball. Although the gorilla does not fit in a basketball game, it is not relevant to the game or to the participant's task to count the number of passes. In the case of the radio dispatchers, they have developed individual and collective strategies to avoid missing important changes. An example of this is the "tagging list" which they used especially after they have been interrupted in order to check updates of incidents graded as immediate. They ignored updates to incidents with a lesser grade. Even though the CDM did not identify change blindness scenarios, it allowed the researcher to elicit strategies used by radio dispatchers to recover from interruptions and handle multiple tasks while monitoring dynamic situations.

Another important point is the use of the mapping tool, or rather the complete lack thereof. Previous research had identified and measured change blindness by evaluating graphical tactical displays. The mapping tool at the BTP control room was not used because it was too cluttered and impractical for their jobs. The radio dispatchers monitored incidents using the command and control text-based system as their main tool. Ignoring a tool that provides vital geographical information might seem unreasonable, however, with appropriate training the command and control text-based system is sufficient.

Although change blindness may have been found if radio dispatchers were asked to use the tactical map, the objective of the field study was to explore change blindness without the constraints of a controlled simulation, not to impose the use of an application that they never employ. While the mapping tool could definitely be improved, there is doubt as to whether it is necessary at all.

Good interface design may have prevented change blindness from occurring at all. The communications and command and control systems used appropriate visual and auditory cues to highlight important events. Because operators are loaded with voice communications, most of the cues are visual. Auditory cues are used only to notify all staff of major incidents and emergencies such as the London bombings in 2005. The interface is message-based and therefore presents discrete changes – an incident is either updated or not updated, a job is either new or acknowledged. This is a different situation from the tactical graphical displays that had been evaluated in previous research. Graphical tactical displays show changes that are continuous where the nature of the state is dependent on multiple components. Missing a change in one of those components could have serious repercussions because the effect on the pattern might not be detected until it is too late. However, the rightness or wrongness of a state in a system like the one used in the BTP control room is far easily determined – each state is binary and it is colour-coded.

The command and control system also provides a thorough logging that gives operators access on demand of the history of events. This logging is colour coded which facilitates the perception of different states within the messages. The 'tagging list' facilitates interruption recovery by allowing radio dispatchers to comprehend the current situation at a glance and shares the same colour-coding rules than the logging system. Thus, the BTP system allows operators to go back not only to where they were before the interruption, but also allows them to check what happened while they were away and rapidly identify a change of status.

Each screen performs one function. For instance, the monitor on the left of the radio dispatcher is for managing the communications system only and the middle monitor is for the command and control system. The fact that different functions are in different screens means that there are visual changes that are specific for each screen. This functionality reduces the possibility of missing certain changes because of window superimposition. An incoming call will only flashed on the communication system, and an incident update will only be presented on the command and control system.

Although this was only an early, exploratory field study, it brings into doubt assumptions about the frequency of change blindness in operational environments. Although some have suggested that change blindness affects operators' situation awareness, this research suggests that change and inattentional blindness might not be operational problems. It seems that operators at the BTP control room have learned to detect important changes at the expense of missing others that do not have an impact on their job performance. If a similar behaviour is found in other command and control rooms, then the assumed causal link between these phenomena and operational error might not be such a severe issue, but further research is required.

The results of this study remind us that these cognitive phenomena are just a consequence of how our attention works and they represent the cost of "our exceptional – and exceptionally useful – ability to focus our minds" (Chabris *et al.*, 2010). In fact, future research should examine the strategies that operators develop to deal with our cognitive limitations, and not only the errors caused by the cognitive limitation.

8.3. Final thoughts

This research has contributed to the field of human-computer interaction by exploring the utility of an innovative display technology for change detection tasks and the occurrence of change blindness in an operational environment.

When evaluating new technology, it is important to analyse not only its capabilities but also be aware of its limitations and the ways users interact with it. As technology continues to progress, new, innovative and speculative devices will be developed, and finding a method to evaluate them is important for the field of human-computer interaction. This research questions claims made for new technologies in areas where lab-based studies suggest using them in operational environments but field work evaluations have not been carried out. It suggests that iteration between the lab and the field, although difficult, could provide valuable insight not only about the technology but also about the operational setting that it could be used in.

One of the objectives of this research was to provide a tool to enhance change detection in operational environments. Although the concept of the multi-layer technology holds promise as a helpful tool, we first have to acknowledge that there are limits to our cognition that technological innovation can not entirely solve. The phenomena of change and inattentional blindness are probably not as severe an issue in operational environments as previous research has argued. Further research is required and the results of this research cannot be generalized until a comparison is made with environments that use tactical graphical displays and others with message-based systems.

This research contributed with interesting results about the depth of the MLD as a tool to enhance change detection. In particular, the experimental design for Experiments 3 and 4 has been innovative. The use of depth to allow direct comparison to detect changes has not been done before. The results suggest that translation and size differences are detected faster and more accurately on the MLD. Although the MLD might not be ready to be used in an operational environment, its depth had a positive effect on change detection tasks. The lessons learnt about the use of depth of the MLD could probably be translated to other three-dimensional displays. Finally, even though the connotations of the word blindness suggest that change and inattentional blindness are severe impediments to human performance, in reality, we should not forget that the ability to focus our attention allows us to avoid distractions and use our limited resources more effectively.

REFERENCES

Aboelsaadat, W., & Balakrishnan, R. (2004, September 6 - 10). *An Empirical Comparison of Transparency on One and Two Layer Displays.* Paper presented at the HCI 2004 – the British HCI Conference.

Allison, R. S., Gillam, B. J., & Vecellio, E. (2009). Binocular depth discrimination and estimation beyond interaction space. *Journal of Vision, 9*(1), 1-14.

Ankrum, D. R. (1999). Visual Ergonomics in the office. *Occupational Health and Safety, 68*(7), 64-74.

Archambault, A., O'Donnell, C., & Schyns, P. (1999). Blind to Object Changes: When Learning the Same Object at Different Levels of Categorization Modifies its Perception. *Psychological Science, 10*(3), 249-255.

Austen, E., & Enns, J. (2000). Change detection: paying attention to detail. *PSYCHE, 6*(11).

Baguley, T. (2004). Introduction to Sphericity. Retrieved October 6, 2011, from http://homepages.gold.ac.uk/aphome/spheric.html#2%29

Beck, D. M., Rees, G., Frith, C. D., & Lavie, N. (2001). Neural Correlates of Change Detection and Change Blindness. *Nature Neuroscience, 4*, 645 - 650

Belke, E., & Meyer, A. S. (2002). Tracking the time course of multidimensional stimulus discrimination: Analyses of viewing patterns and processing times during "same"-"different" decisions. *European Journal of Cognitive Psychology, 14*(2), 237-266.

Bell, G., Craig, R., & Simmiss, T. (2007). Moire interference in multi layered displays. *Journal of the Society for Information Display, 15*(11), 883-888.

Bell, G., Engel, G., Searle, M. J., & Evanicky, D. (2006). USA Patent No.

Bell, G. P., Craig, R., Paxton, R., Wong, G., & Galbraith, D. (2008). *Beyond Flat Panels - Multi Layer Displays with Real Depth*. Paper presented at the Conference Name|. Retrieved Access Date|. from URL|.

Bishop, C. (2005). *Usability Issues of Multiple-Layer Display Technology.* University of Canterbury, Christchurch.

Boduroglu, A., & Shah, P. (2009). Effects of spatial configurations on visual change detection: An account of bias changes. *Memory & Cognition, 37*(8), 1120-1131.

Bolia, R., Todd, N., & Vidulich, M. (2004). A multi-layer visual display for air battle managers: Effects of depth and transparency on performance and workload in a dual-task scenario (Abstract). *Human factors and Aerospace Safety, 4*(3), 181 - 193.

Boot, W. R., Becic, E., & Kramer, A. F. (2007). Temporal limitations in multiple target detection in a dynamic monitoring task. *Human Factors, 49*(5), 897-906.

Boot, W. R., Becic, E., & Kramer, A. F. (2009). Stable individual differences in search strategy?: The effect of task demands and motivational factors on scanning strategy in visual search. *Journal of Vision, 9*(3).

Bremmer, F., Kubischik, M., Hoffmann, K. P., & Krekelberg, B. (2009). Neural Dynamics of Saccadic Suppression. *Journal of Neuroscience, 29*(40), 12374-12383.

Bridgeman, B., Vanderheijden, A. H. C., & Velichkovsky, B. M. (1994). A Theory of Visual-Stability across Saccadic Eye-Movements. *Behavioral and Brain Sciences, 17*(2), 247-258.

Bruce, V., Green, P. R., & Georgeson, M. A. (2003). *Visual Perception: Physiology, Psychology, & Ecology*. New York: Psychology Press.

Brunel, N., & Ninio, J. (1997). Time to detect the difference between two images presented side by side. *Cognitive Brain Research, 5*(4), 273-282.

Bruzzone, L., Prieto, D. F., & Serpico, S. B. (1999). A neural-statistical approach to multitemporal and multisource remote-sensing image classification. *Ieee Transactions on Geoscience and Remote Sensing, 37*(3), 1350-1359.

Bunks, C. (2000). Section 5.6 The Blending Modes Retrieved October 17, 2009, from http://gimp-savvy.com/BOOK/index.html?node55.html

Burgess-Limerick, R., Plooy, A., Fraser, K., & Ankrum, D. R. (1999). The influence of computer monitor height on head and neck posture. *International Journal of Industrial Ergonomics, 23*(3), 171-179.

BusinessWire. (2009). PureDepth(TM) MLD(TM) Technology Drives Japanese Pachislot Game "Onihama" to Massive Popularity [Electronic Version]. *Business WIre*. Retrieved October 13, 2009 from http://ir.puredepth.com/releasedetail.cfm?ReleaseID=378473&l=en.

Card, S. K., Moran, T. P., & Newell, A. (1983). *The psychology of human-computer interaction* Hillsdale, NJ: Lawrence Erlbaum Associates.

Carr, S. M., Meehan, J. W., & Phillips, J. G. (2006). Does a Multi-Layer Display Reduce the Effect of Non-Target Flankers. *Human Factors and Ergonomics Society Annual Meeting Proceedings, 50*, 1647-1651.

Carrasco, M., Evert, D., Chang, I., & Kantz, S. (1995). The eccentricity effect: target eccentricity affect performance on conjunction searches. *Perception and Psychophysics, 57*(8), 1241-1261.

Chabris, C., & Simons, D. (2010). *The Invisible Gorilla: And Other Ways Our Intuition Deceives Us*. London: HarperCollins Publishers Limited.

Chadwick, E. (1999). An Artist's Real-Time 3D Glossary. Retrieved October 17, 2009, from http://www.ericchadwick.com/portfolio/glossary/glossary.html#B

Chan, A. H. S., & Ng, A. W. Y. (2009). Perceptions of implied hazard for visual and auditory alerting signals. *Safety Science, 47*(3), 346-352.

Chan, A. H. S., & Tang, N. Y. W. (2007). Visual lobe shape and search performance for targets of different difficulty. *Ergonomics, 50*(2), 289-318.

Chan, H. S., & Courtney, A. J. (1997). Visual performance on synchronous and asynchronous peripheral target detection tasks. *Human Factors and Ergonomics in Manufacturing, 7*(3), 197-210.

Changingminds.org. (2010). Cramer's V. Retrieved November 7, 2010, from http://changingminds.org/explanations/research/analysis/cramers_v.htm

Chastain, G., & Cheal, M. (1999). Attention Effects of Abrupt-Onset Precues with Central, Single-Element, and Multiple-Element Precues. *Consciousness and Cognition, 8*(4), 510-528.

Cooperstock, J. R., & Wang, G. (2009, January 19). *Stereoscopic display technologies, interaction paradigms, and rendering approaches for neurosurgical visualization.* Paper presented at the Stereoscopic Displays and Applications XX San Jose, CA.

Coutant, B. E., & Westheimer, G. (1993). Population distribution of stereoscopic ability. *Ophthalmic and Physiological Optics, 13*(1), 3-7.

Curran, T., Gibson, L., Horne, J. H., Young, B., & Bozell, A. P. (2009). Expert image analysts show enhanced visual processing in change detection. *Psychonomic Bulletin & Review, 16*(2), 390-397.

De la Rosa, S., Moraglia, G., & Schneider, B. A. (2008). The Magnitude of Binocular Disparity Modulates Search Time for Targets Defined by a Conjunction of Depth and Colour. *Canadian Journal of Experimental Psychology-Revue Canadienne De Psychologie Experimentale, 62*(3), 150-155.

Deubel, H., Schneider, W. X., & Bridgeman, B. (2002). Transsaccadic memory of position and form. *Brain's Eye: Neurobiological and Clinical Aspects of Oculomotor Research, 140*, 165-180.

Di Vita, J., Nugent, W., Obermayer, R., & Linville, J. (2004). Verification of Change Blindness Phenomenon while Managing Critical Events on a Combat Information Display (Attentional Processes). *Human Factors, 46*(2), 205-218.

Dimmick, F. L. (1944). The retina. The anatomy and the histology of the retina in man, ape, and monkey, including the consideration of visual functions, the history of physiological optics, and the histological laboratory technique.[Review]. *Psychological Bulletin, 41*(3), 192-194.

Diner, D. B., & Fender, D. H. (1993). *Human Engineering in Stereoscopic Display Devices*. New York, NY: Plenum Press.

Dodgson, N. A. (2004). Variation and extrema of human interpupillary distance. *Stereoscopic Displays and Virtual Reality Systems Xi, 5291*, 36-46.

Dornhoefer, S. M., Unema, P. J. A., & Velichkovsky, B. M. (2002). Blinks, blanks and saccades: how blind we really are for relevant visual events. *Brain's Eye: Neurobiological and Clinical Aspects of Oculomotor Research, 140*, 119-131.

Duchowski, A. (2000). *Eye-Based Interaction in Graphical Systems: Theory & Practice*. Paper presented at the Conference Name|. Retrieved Access Date|. from URL|.

Dunser, A., Billinghurst, M., & Mancero, G. (2008, December 8 -12). *Evaluating visual search performance with a multi layer display.* Paper presented at the 20th Australasian Conference on Computer-Human Interaction: Designing for Habitus and Habitat, Cairns, Australia.

Dunser, A., & Mancero, G. (2009). The use of Depth in Change Detection and Multiple Object Tracking. *The Ergonomics Open Journal, 2*, 142-149.

Durlach, P. (2004a). *Army Digital Systems and Vulnerability to Change Blindness* (No. 0704-0188). Orlando, FL: U.S. Army Research Institute for the Behavioral and Social Sciences.

Durlach, P. (2004b). Change Blindness and its Implications for Complex and Control Systems Design and Operator Training. *Human Computer Interaction, 19*, 423-451.

Durlach, P. J., Kring, J. P., & Bowens, L. D. (2008). Detection of icon appearance and disappearance on a digital situation awareness display. *Military Psychology, 20*(2), 81-94.

Ellis, S., R. Tyler, M., Kim, W., S., MvGreevy, M., W., & Stark, L. (1985). *Visual enhancements for perspective displays: perspective parameters.* Paper presented at the IEEE Conference on Cybernetics and Society.

Engbert, R., & Kliegl, R. (2003). Microsaccades uncover the orientation of covert attention. *Vision Research, 43*(9), 1035-1045.

Ericsson, K. A., Krampe, R. T., & Tesch-ROmer, C. (1993). The Role of Deliberate Practice in the Acquisition of Expert Performance. *Psychological Review, 100*(3), 363-406.

Eysenck, M. W., & Keane, M. T. (2005). *Cognitive psychology: a student's handbook* (Fifth Edition ed.). New York, NY: Psychology Press.

Farell, B. (1985). Same Different Judgments - a Review of Current Controversies in Perceptual Comparisons. *Psychological Bulletin, 98*(3), 419-456.

Field, A. (2009). *Discovering Statistics Using SPSS* (Third ed.). London: SAGE Publications.

Fisher, D. L., Coury, B. G., Tengs, T. O., & Duffy, S. A. (1989). Minimizing the Time to Search Visual Displays: The Role of Highlighting. *Human Factors, 31*(2), 167 - 182.

Fisher, D. L., & Tan, K. C. (1989). Visual Displays: The Highlighting Paradox. *Human Factors, 31*(1), 17-30.

Froner, B., Holliman, N. S., & Liversedge, S. P. (2008). A comparative study of fine depth perception on two-view 3D displays. *Displays, 29*(5), 440-450.

Galster, S. M., Bolia, R., & Brown, R. D. (2006). The efficacy of advanced visual display technologies in simulated airborne command and control environments. In D. d. Waard, K. A. Brookhuis & A. Toffetti (Eds.), *Developments in Human Factors in Transportation, Design and Evaluation* (pp. 1-10). Maastricht, The Netherlands: Shaker Publishing.

Gerber, H. (2005). Missing Values Analysis and Imputation Retrieved November 12, 2010, from www.itc.virginia.edu/research/talks/missing.ppt

Girden, E. R. (1992). *ANOVA: repeated measures*: Sage Publications.

Glenstrup, A. J., & Engell-Nielsen, T. (1995). Eye Movements. Retrieved September 25, 2009, from http://www.diku.dk/hjemmesider/ansatte/panic/eyegaze/node16.html

Halle, M. W. (1997). Autostereoscopic Displays and Computer Graphics. *Computer Graphics, ACM SIGGRAPH, 31*(2), 58-62.

Hayes, J., Moore, A., & Wong, B. L. W. (2006). *Information Layering to declutter displays for Emergency Ambulance Dispatch*. Paper presented at the Conference Name|. Retrieved Access Date|. from URL|.

Healey, C. G. (2005). Perception in Visualization. Retrieved November 4, 2011, from http://www.csc.ncsu.edu/faculty/healey/PP/index.html

Henderson, J. M., & Hollingworth, A. (1999). The role of fixation position in detecting scene changes across saccades. *Psychological Science, 10*(5), 438-443.

Herslund, M. B., & Jorgensen, N. O. (2003). Looked-but-failed-to-see-errors in traffic. *Accident Analysis and Prevention, 35*(6), 885-891.

Hill, T., & Lewicki, P. (2005). STATSOFT Electronic Statistics Book. Retrieved November 7, 2010, from http://www.statsoft.com/textbook/

Holliday, I. E., & Braddick, O. J. (1991). Pre-attentive detection of a target defined by stereoscopic slant. *Perception, 20*, 355 - 362.

Holliman, N. (2004). *Mapping perceived depth to regions of interest in stereoscopic images*. Paper presented at the Conference Name|. Retrieved Access Date|. from URL|.

Holliman, N. (2006). Three-dimensional display systems. In J. Dakin & R. G. W. Brown (Eds.), *Handbook of Optoelectronics* (Vol. 1, pp. 1067 - 1100). Boca Raton, FL: Taylor & Francis Group.

Howard, I. P., & Rogers, B. J. (1995). *Binocular vision and stereopsis*. New York: Oxford University Press.

Howell, D. C. (2009a). Multiple Comparisons with Repeated Measures [Electronic Version]. Retrieved February 17 2009 from http://www.uvm.edu/~dhowell/StatPages/More_Stuff/RepMeasMultComp/RepMeasMultComp.html.

Howell, D. C. (2009b, June 3, 2009). Treatment of Missing Data. Retrieved August 9, 2009, from http://www.uvm.edu/~dhowell/StatPages/More_Stuff/Missing_Data/Missing.html

Hutchins, S., Pirolli, P. L., & Card, S. K. (2004). *A New Perspective on Use of the Critical Decision Method with Intelligence Analysts*. Monterey, CA: Naval Postgraduate School, Information Science Department.

Hyun, J. S., Woodman, G. F., Vogel, E. K., Hollingworth, A., & Luck, S. J. (2009). The Comparison of Visual Working Memory Representations With Perceptual Inputs. *Journal of Experimental Psychology-Human Perception and Performance, 35*(4), 1140-1160.

Irwin, D. E. (1991). Information integration across saccadic eye movements. *Cognitive Psychology, 23*(3), 420-456.

Irwin, D. E., & Gordon, R. (1998). Eye Movements, Attention and Trans-saccadic Memory. *Visual Cognition, 5*(1), 127 - 155.

Iwasaki, M., & Inomata, H. (1986). Relation between Superficial Capillaries and Foveal Structures in the Human Retina. *Investigative Ophthalmology & Visual Science, 27*(12), 1698-1705.

Jonides, J., Irwin, D. E., & Yantis, S. (1982). Integrating Visual Information from Successive Fixations. *Science, 215*(4529), 192-194.

Kieras, D. (2001). *Using the Keystroke-Level Model to Estimate Execution Times*. Michigan: University of Michigan.

Klein, G., Calderwood, R., & Macgregor, D. (1989). Critical decision method for eliciting knowledge. *IEEE Transactions on Systems, Man and Cybernetics, 19*, 462–472.

Kosara, R., Miksch, S., & Hauser, H. (2002). Focus+Context Taken Literally. *IEEE Computer Graphics & Applications (CG&A), Special Issue on Information Visualization, 22*(1), 22-29.

Kroft, P., & Wickens, C. (2003). Displaying multi-domain graphical database information. *Information Design Journal, 11*(1), 44 - 52.

Kuhn, G., & Findlay, J. M. (2010). Misdirection, attention and awareness: Inattentional blindness reveals temporal relationship between eye movements and visual awareness. *Quarterly Journal of Experimental Psychology, 63*(1), 136-146.

Lansdown, J. (1996). *Visual Perception*: Middlesex University - Centre for Electronic Arts.

Levin, D. T. (2002). Change blindness blindness as visual metacognition. *Journal of Consciousness Studies, 9*(5-6), 111-130.

Levin, D. T., Drivdahl, S. B., Momen, N., & Beck, M. R. (2002). False predictions about the detectability of visual changes: The role of beliefs about attention, memory, and the continuity of attended objects in causing change blindness blindness. *Consciousness and Cognition, 11*(4), 507-527.

Linde, I. V. D. (2003). *Space-variant Perceptual Image Compression for Gaze-contingent Stereoscopic Displays.* Anglia Polytechnic University, UK.

Little, R. J. A., & Rubin, D. B. (1987). *Statistical Analysis with Missing Data*. New York: John Wiley and Sons.

Lu, D., Mausel, P., BrondÃ-zio, E., & Moran, E. (2004). Change detection techniques. *International Journal of Remote Sensing, 25*(12), 2365-2401.

Mack, A. (2007). Hochberg and Inattentional Blindness. In M. A. Peterson, B. Gillam & H. A. Sedgwick (Eds.), *In the Mind's Eye: Julian Hochberg on the perception of pictures, films, and the world* (pp. 634). New York: Oxford University Press.

Mack, A., & Rock, I. (1999). Inattentional Blindness: An Overview. *Psyche, 5*(3).

Malin, J. T., Schreckenghost, D. L., Woods, D. D., Potter, S. S., Johannesen, L., Holloway, M.*, et al.* (1991). *Making intelligent systems team players: Case studies and design issues. Volume 1: Human-computer interaction design* (No. NAS 1.15:104738; NASA-TM-104738-VOL-1; S-643). Houston, TX: NASA.

Mancero, G., & Wong, B. L. W. (2008). *An Evaluation of Perceptual Depth to Enhance Change Detection.* Paper presented at the Human Factors and Ergonomics Society 52nd Annual Meeting

Mancero, G., Wong, B. L. W., & Amaldi, P. (2008). *The Utility of Depth in Multi-Layered Displays for Supporting Change Detection* (No. 2008-4-001). London: Middlesex University and European Office of Aerospace Research and Development.

Mancero, G., Wong, B. L. W., & Loomes, M. (2009a, November 23-27). *Radio Dispatcher's Interruption Recovery Strategies* Paper presented at the OzCHI 2009, Melbourne,Australia.

Mancero, G., Wong, W., & Loomes, M. (2009b, November 30 - October 2). *Change Blindness and Situation Awareness in a Police C2 Environment.* Paper presented at the European Conference on Cognitive Ergonomics (ECCE), Helsinki, Finland.

Masoodian, M., McKoy, S., Rogers, W. J., & Ware, D. (2004). *DeepDocument: use of a multi-layered display to provide context awareness in text editing*. Hamilton, NZ: University of Waikato, Department of Computer Science.

Mauchly, J. W. (1940). Significance Test for Sphericity of a Normal n-Variate Distribution. *The Annals of Mathematical Statistics*, 204 -209.

McConkie, G. W., & Currie, C. B. (1996). Visual stability across saccades while viewing complex pictures. *Journal of Experimental Psychology-Human Perception and Performance, 22*(3), 563-581.

McKnight, P. E., McKnight, K. M., Sidani, S., & Figueredo, A. J. (2007). *Missing data: a gentle introduction*. New York: The Guildford Press.

Meiran, N., & Marciano, H. (2002). Limitations in advance task preparation: Switching the relevant stimulus dimension in speeded same-different comparisons. *Memory & Cognition, 30*(4), 540-550.

Militello, L., Wong, B. L. W., Kirschenbaum, S., & Patterson, E. (2009). Chapter 18. Systematizing Discovery in Cognitive Task Analysis. . In K. L. Mosier & U. M. Fischer (Eds.), *Informed by Knowledge: Expert Performance in Complex Situations (In Press)*.

Most, S. B., & Astur, R. S. (2007). Feature-based attentional set as a cause of traffic accidents. *Visual Cognition, 15*(2), 125-132.

Most, S. B., Scholl, B. J., Simons, D. J., & Clifford, E. R. (2005). What You See is What You Set: Sustained Inattentional Blindness and the Capture of Awareness. *Psychological Review, 112*(1), 217 - 242.

Most, S. B., Simons, D. J., Scholl, B. J., & Chabris, C. (2000). Sustained Inattentional Blindness: The Role of Location in the Detection of Unexpected Dynamic Events *PSYCHE, 6*(14).

Most, S. B., Simons, D. J., Scholl, B. J., Jimenez, R., Clifford, E., & Chabris, C. F. (2001). How not to be seen: The contribution of similarity and selective ignoring to sustained inattentional blindness. *Psychological Science, 12*(1), 9-17.

Mulckhuyse, M., van Zoest, W., & Theeuwes, J. (2008). Capture of the eyes by relevant and irrelevant onsets. *Experimental Brain Research, 186*(2), 225-235.

Muller, H. J., & Rabbitt, P. M. A. (1989). Reflexive and Voluntary Orienting of Visual-Attention - Time Course of Activation and Resistance to Interruption. *Journal of Experimental Psychology-Human Perception and Performance, 15*(2), 315-330.

Muthard, E. K., & Wickens, C. D. (2002). Change Detection after Preliminary Flight Decisions: Linking Planning Errors to Biases in Plan Monitoring. *Human Factors and Ergonomics Society Annual Meeting Proceedings, 46*, 91-95.

Muthard, E. K., & Wickens, C. D. (2003). *Factors that Mediate Flight Plan Monitoring and Errors in Plan Revision Planning under Automated and High Workload Conditions.* Paper presented at the 12th International Symposium on Aviation Psychology, Dayton, OH.

Naikar, N. (1998). *Perspective Displays: A review of Human Factors Issues* (Technical Report No. MI/9/436). Melbourne: Aeronautical and Maritime Research Laboratory.

Nakayama, K., & Silverman, G. H. (1986). Serial and parallel processing of visual feature conjunctions. *Nature, 320*(6059), 264-265.

Nave, R. (2001). HyperPhysics. Retrieved October 31, 2009, from http://hyperphysics.phy-astr.gsu.edu/hbase/vision/retina.html#c1

Neisser, U. (1976). *Cognition and reality : principles and implications of cognitive psychology*. San Francisco: W.H. Freeman.

Neisser, U., & Becklen, R. (1975). Selective Looking - Attending to Visually Specified Events. *Cognitive Psychology, 7*(4), 480-494.

Nikolic, M. I., Orr, J. M., & Sarter, N. B. (2004). Why Pilots Miss the Green Box: How Display Context Undermines Attention Capture. *The International Journal of Aviation Psychology, 14*(1), 39-52.

Nikolic, M. I., & Sarter, N. B. (2001). Peripheral visual feedback: A powerful means of supporting effective attention allocation in event-driven, data-rich environments. *Abstract*, *Human Factors* (Vol. 43, pp. 30-38).

Ninio, J. (1998). Acquisition of shape information in working memory, as a function of viewing time and number of consecutive images: evidence for a succession of discrete storage classes. *Cognitive Brain Research, 7*(1), 57-69.

Ninio, J. (2004). Testing sequence effects in visual memory: clues for a structural model. *Acta Psychologica, 116*(3), 263-283.

O'Regan, J. K., Deubel, H., Clark, J. J., & Rensink, R. A. (2000). Picture changes during blinks: Looking without seeing and seeing without looking. *Visual Cognition, 7*(1-3), 191-211.

O'Regan, J. K., & Noe, A. (2001). A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences, 24*(5), 939-+.

O'Toole, A. J., & Walker, C. L. (1997). On the preattentive accessibility of stereoscopic disparity: Evidence from visual search. *Perception and Psychophysics, 59*(2), 202 -218.

Obermayer, R. W., & Nugent, W. A. (1999). Human-Computer Interaction for Alert Warning and Attention Allocation Systems of the Multi-Modal Watchstation. Retrieved February 19 2006, from http://www.manningaffordability.com/S&tweb/PUBS/HCI-AlertWarning-MMWS/HCI-AlertWarning-MMWS.htm

Owens, D. A., & Wolf-Kelly, K. (1987). Near Work, Visual Fatigue, and Variations of Oculomotor Tonus. *Investigative Ophthalmology and Visual Science, 28*, 743-749.

Palmer, E. M., Clausner, T. C., & Kellman, P. J. (2008). Enhancing air traffic displays via perceptual cues. *ACM Trans. Appl. Percept., 5*(1), 1-22.

Palmer, J., Verghese, P., & Pavel, M. (2000). The psychophysics of visual search. *Vision Research, 40*(10 -12), Pages 1227-1268.

Parkin, A. J. (2000). *Essential cognitive psychology*. Hove, East Sussex: Psychology Press.

Patterson, R., Becker, S., Boucek, G. S., & Phinney, R. (1994). Depth perception in stereoscopic displays. *Journal of the Society for Information Display, 2*(2), 105-112.

Pessoa, L., & Ungerleider, L. (2004). Neural Correlates of Change Detection and Change Blindness in a Working Memory Task *Cerebral Cortex, 14*, 511-520.

Phillips, W. A. (1974). On the distinction between sensory storage and short-term visual memory [Abstract]. *Perception and Psychophysics, 16*, 283-290.

Pirolli, P., & Card, S. (1999). Information foraging. *Psychological Review, 106*(4), 643-675.

Podczerwinski, E. S., Wickens, C. D., & Alexander, A. L. (2002). *Exploring the "Out of Sight, Out of Mind" Phenomenon in Dynamic Settings Across Electronic Map Displays* (Technical

Report No. ARL-01-8/NASA-01-4). Moffett Field, CA: Aviation Research Lab Institute of Aviation.

Polyak, S. (1941). *The Retina*. Chicago: University of Chicago Press.

Prema, V., Roberts, G., & Wuensche, B. C. (2006, 27-29 November). *3D Visualisation Techniques for Multi-Layer Display Technology.* Paper presented at the IVCNZ '06, Great Barrier Island, New Zealand.

Pringle, H. l., Irwin, D. E., Kramer, A. F., & Atchley, P. (2001). The role of attentional breadth in perceptual change detection. *Psychonomic Bulletin & Review, 8*(1), 89-95.

PRNewswire. (2007). IGT and PureDepth Launch World's First MLD-Based Product [Electronic Version]. *COMTEX News Network*. Retrieved October 13 from http://ir.puredepth.com/releasedetail.cfm?ReleaseID=272454&l=en.

PureDepth. (2007). *White Paper Multi-Layer Displays (MLDs)* (Report). Auckland Pure Depth.

Pylyshyn, Z. (2006). *Seeing and Visualizing: It's not what you think*. Cambridge, MA: Massachussetts Institute of Technology.

Quinlan, P. T. (2003). Visual feature integration theory: Past, present, and future. *Psychological Bulletin, 129*(5), 643-673.

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology, 62*(8), 1457-1506.

Remington, R. W., Johnston, J. C., Ruthruff, E., Gold, M., & Romera, M. (2000). Visual Search in Complex Displays: Factors Affecting Conflict Detection by Air Traffic Controllers. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 42*(3), 349-366.

Rensink, R. (2000). Visual Search for Change: A Probe into the Nature of Attentional Processing In D. J. Simons (Ed.), *Change Blindness and Visual Memory* (pp. 345 - 377): Psychology Press.

Rensink, R. (2002). Change Detection. *Annual Review Psychology, 53*, 245-277.

Rensink, R., O'Regan, J. K., & Clark, J. (1997). To See or not to See: The Need for Attention to Perceive Changes in the Scene. *Psychological Science, 8*, 368-373.

Rensink, R., O'Regan, J. K., & Clark, J. (1999). Change-blindness as a result of mudsplashes. *Nature, 398(6722)*, 34.

Richard, C. M., Wright, R. D., Ee, C., Prime, S. L., Shimizu, Y., & Vavrik, J. (2002). Effect of a Concurrent Auditory Task on Visual Search Performance in a Driving-Related Image-Flicker Task. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 44*(1), 108-119.

Rosielle, L. J., & Scaggs, W. J. (2008). What if they knocked down the library and nobody noticed? The failure to detect large changes to familiar scenes. *Memory, 16*(2), 115 - 124.

Ross, J., Morrone, M. C., Goldberg, M. E., & Burr, D. C. (2001). Changes in visual perception at the time of saccades. *Trends in Neurosciences, 24*(2), 113-121.

Scholl, Noles, Pasheva, & Sussman. (2003). Talking on a cellular telephone dramatically increases 'sustained inattentional blindness'. *Journal of Vision, 3*(9), 156-156.

Schreij, D., Owens, C., & Theeuwes, J. (2008). Abrupt onsets capture attention independent of top-down control settings. *Perception & Psychophysics, 70*(2), 208-218.

Scott-Brown, K. C., Baker, M. R., & Orbach, H. (2000). Comparison Blindness. *Visual Cognition, 7*(1/2/3), 253-267.

Simons, D. J. (2007). Inattentional blindness [Electronic Version]. *Scholarpedia*, 2, 3244. Retrieved July 19 from http://www.scholarpedia.org/article/Inattentional_blindness.

Simons, D. J., & Ambinder, M. S. (2005). Change blindness - Theory and consequences. *Current Directions in Psychological Science, 14*(1), 44-48.

Simons, D. J., & Chabris, C. F. (1999). Gorillas in the midst: Sustained Inattentional Blindness for Dynamic Events. *Perception, 28*(9), 1059-1074.

Simons, D. J., Franconeri, S., & Reimer, R. (2000). Change Blindness in the absence of a visual disruption. *Perception, 29*, 1143-1154.

Simons, D. J., & Levin, D. (1997a). Change Blindness. *Trends in Cognitive Sciences, 1*(7), 261-267.

Simons, D. J., & Levin, D. (1997b). Failure to Detect Changes to Attended Objects in Motion Pictures. *Psychonomic Bulletin and Review, 4*, 501-506.

Simons, D. J., & Levin, D. (1998). Failure to Detect Changes to People in Real-World Interactions. *Psychonomic Bulletin and Review, 5*(644-649).

Simons, D. J., & Rensink, R. A. (2005). Change blindness: past, present, and future. *Trends in Cognitive Sciences, 9*(1), 16-20.

Singh, D. (2005). *Software Development for Multi-Layer Displays*. Auckland: Pure Depth.

Smallman, H. S., & St. John, M. (2003). *CHEX (Change History EXplicit): New HCI concepts for change awareness.* Paper presented at the 46th Annual Meeting of the Human Factors and Ergonomics Society, Santa Monica, CA.

St. John, M., & Cowen, M. B. (1999). *Use of Perspective View Displays for Operational Tasks*. San Diego: Pacific Science and Engineering Group and SSC San Diego.

St. John, M., Smallman, H. S., Boynton, R. M., Oonk, H. M., & Cowen, M. B. (2000). *Navigating two-dimensional and perspective views of terrain.* (Technical No. 1827). San Diego, CA: SPAWAR Systems Center.

St. John, M., Smallman, H. S., & Manes, D. I. (2005). *Recovery from interruptions to a dynamic monitoring task: the beguiling utility of instant replay*. Paper presented at the Conference Name|. Retrieved Access Date|. from URL|.

St. John, M., Smallman, H. S., & Manes, D. I. (2007). *Interruption Recovery Tools for Team Collaboration* Paper presented at the Human Factors and Ergonomics 51st Annual Meeting, Baltimore, Maryland.

Stanton, N., Salmon, P., Walker, G., Baber, C., & Jenkins, D. P. (2005). *Human Factors Methods: A Practical Guide for Engineering and Design*. Aldershot, England: Ashgate Publishing Ltd.

Straker, L. M. (2001). Visual Display Units: Positioning for human use. In W. Karwowski (Ed.), *International encyclopedia of ergonomics and human factors* (Vol. 2, pp. 975 - 977). London: Taylor & Francis.

Tabachnick, B. G., & Fidell, L. S. (1996). *Using Multivariate Statistics*. New York: Harper Collins College Publishers.

Thackray, R. I., & Touchstone, R. M. (1991). Effects of Monitoring under High and Low Taskload on Detection of Flashing and Colored Radar Targets. *Ergonomics, 34*(8), 1065-1081.

Theeuwes, J. (1994). Endogenous and Exogenous Control of Visual Selection. *Perception, 23*(4), 429-440.

Theeuwes, J. (1995). Abrupt Luminance Change Pops out - Abrupt Color-Change Does Not. *Perception & Psychophysics, 57*(5), 637-644.

Theeuwes, J., Atchley, P., & Kramer, A. F. (1998). Attentional control within 3-D space. *Journal of Experimental Psychology: Human Perception and Performance, 24*(5), 1476–1485.

Treisman, A. (1982). Perceptual grouping and attention in visual search for features and for objects. . *Exp Psychol Hum Percept Perform, 8*(2), 194-214.

Treisman, A. (1985). Preattentive Processing in Vision. *Computer Vision, Graphics, and Image Processing, 31*(2), 156-177.

Treisman, A., & Gelade, G. (1980). A Feature-Integration Theory of Attention. *Cognitive Psychology, 12*, 97-136.

UPA. (2009). What is User-Centered Design? Retrieved October 12, 2009, from http://www.usabilityprofessionals.org/usability_resources/about_usability/what_is_ucd.html

Varakin, D. A., Levin, D., & Fidler, R. (2004). Unseen and Unaware: Implications of Recent Research on Failures of Visual Awareness for Human-Computer Interface Design. *Human-Computer Interaction, 19*, 389-422.

Vredenburg, K., Isensee, S., & RIghi, C. (2002). *User-Centered Design: An Integrated Approach*. Upper Saddle River, NJ: Prentice Hall PTR.

Werner, S., & Thies, B. (2000). Is Change Blindness Attenuated by Domain-specific Expertise? An Expert-Novices Comparison in Change Detection in Football Images *Visual Cognition, 7*, 163-173.

Wickens, C. D., & Alexander, A. L. (2009). Attentional Tunneling and Task Management in Synthetic Vision Displays. *International Journal of Aviation Psychology, 19*(2), 182-199.

Wickens, C. D., Ambinder, M., & Alexander, A. (2004). *The Role of Highlighting in VIsual Search Through Maps.* Paper presented at the Proceedings of the Human Factors and Ergonomics Society 48th Annual Meeting, Santa Monica.

Wickens, C. D., & Haskell, I. D. (1993). Two and Three Dimensional Displays for Aviation: A Theoretical and Empirical Comparison. *The International Journal of Aviation Psychology, 3*(2), 87-109.

Wickens, C. D., & McCarley, J. S. (2008). *Applied Attention Theory*. Boca Raton: CRC Press.

Wickens, C. D., Muthard, E. K., Alexander, A. L., VanOlffen, P., & Podczerwinski, E. (2003). *The Influences of Display Highlighting and Size And Event Eccentricity for Aviation Surveillance.*

Paper presented at the 47th Annual Meeting of the Human Factors and Ergonomics Society, Santa Monica, CA.

Wickens, C. D., Todd, S., & Seidler, K. (1989). *Three-dimensional displays: Perception, Implementation and Applications*. Savoy, IL: Crew System Ergonomics Information Analysis Centre.

Wiley, R. (2008). OLED TV vs. LCD TV. Retrieved November 7, 2010, from http://www.oledbuyingguide.com/oled-tv-articles/oled-tv-vs-lcd-tv.html

Wolfe, J. M. (2007). Guided Search 4.0: Current Progress with a Model of Visual Search. In W. Gray (Ed.), *Integrated Models of Cognitive Systems* (pp. 99 - 119). New York: : Oxford.

Wolfe, J. M., & Horowitz, T. S. (2004). What attributes guide the deployment of visual attention and how do they do it? *Nat Rev Neurosci, 5*(6), 495-501.

Wong, B. L. W. (2004). Data analysis for the Critical Decision Method. In D. Diaper & N. Stanton (Eds.), *Task Analysis for Human-Computer Interaction* (pp. 327-346). Mahwah, NJ: Lawrence Erlbaum Associates.

Wong, B. L. W. (2006). The Critical Decision Method. In W. Karwowski (Ed.), *International Encyclopedia of Human Factors and Ergonomics* (pp. 3067-3073): CRC Press.

Wong, B. L. W., & Blandford, A. (2002, 25 - 27 November). *Analysing Ambulance Dispatcher Decision Making: Trialing Emergent Themes Analysis.* Paper presented at the Human Factors 2002, the Joint Conference of the Computer Human Interaction Special Interest Group and The Ergonomics Society of Australia, Melbourne, Australia.

Wong, B. L. W., Joyekurun, R., Nees, A., Amaldi, P., & Villanueva, R. (2005). Information Layering, Depth and Transparency Effects on Multi-Layered Displays for Command and Control. *Human Factors and Ergonomics Society Annual Meeting Proceedings, 49*, 352-356.

Wong, B. L. W., Mansour, H., Joyekurun, R., Amaldi, P., Nees, A., & Villanueva, R. (2005, November 23 -25). *Depth, Layering and Transparency: Developing Design Techniques.* Paper presented at the OZCHI 2005, Canberra, Australia.

Yahiro, M., & Bell, G. (2006). New Zealand Patent No.: PureDepth.

Yantis, S. (1998). Control of visual attention. In H. Pashler (Ed.), *Attention* (pp. 223-256). East Sussex, UK: Psychology Press

Yantis, S., & Jonides, J. (1984). Abrupt Visual Onsets and Selective Attention - Evidence from Visual-Search. *Journal of Experimental Psychology-Human Perception and Performance, 10*(5), 601-621.

Yantis, S., & Jonides, J. (1990). Abrupt Visual Onsets and Selective Attention - Voluntary Versus Automatic Allocation. *Journal of Experimental Psychology-Human Perception and Performance, 16*(1), 121-134.

Yeh, M., & Wickens, C. (2000). *Attention Filtering In the Design of Electronic Map Displays: A Comparison of Color-Coding, Intensity Coding, and Decluttering Techniques* (No. Technical Report ARL-00-4/FED-LAB-00-2). Illinois: Aviation Research Lab - Institute of Aviation

APPENDICES

Appendix I. Stereoacuity Test

Opticians use several tests to measure a person's stereoacuity. These tests help to identify vision problems and conduct stereopsis, amblyopia, suppression, and strabismus testing. The most common stereoacuity test is the Frisby Davis Stereo Test which consists of a square transparent plate on which four similar patterns (resembling a random-dot stereogram) are printed on one side. In the central part of one of the four patterns is a circular area, which is printed on the other side of the plate and can appear in depth. The plate (made of plastic or glass) comes in three thicknesses: 6, 3 and 1mm.

Frisby Davis Stereo Test Plates. Source: http://tinyurl.com/yhejl69

By using the three plates and presenting them at different distances the test can produce a retinal disparity of the circular area between 600 and 7 seconds of arc. Opticians expect an adult to have a stereoacuity of 40 arc seconds.

In this test the patient's head must be kept still to avoid monocular cues. The plate can be turned upside down or rotated to alter the position of the pattern with relief.

Another test for stereo depth perception testing is known as the Stereo Fly. The Stereo Fly evaluates both gross and fine stereo vision. This test uses a large image of a house fly and thus its name. The image of the fly is useful for testing children, as they respond to large objects well. The test also features targets and animals for further stereo testing. The test includes a graded circle test (800 to 40 seconds of arc), an animal test for children (400 to 100 seconds of arc) However, the test only works with the use of the stereo glasses.

Stereo Fly Test. Source: http://tinyurl.com/yzg67sb

Appendix II. Ethics Approval – Middlesex University Ethics Committee and the AFRL – Wright Site Institutional Review Board

DEPARTMENT OF THE AIR FORCE
HEADQUARTERS UNITED STATES AIR FORCE
WASHINGTON DC

NOV 7 2006

MEMORANDUM FOR Middlesex University, School of Computing Science
ATTN: Professor Colin Tully,
Director, Research and Postgraduate Studies

FROM: HQ USAF/SGRC
5201 Leesburg Pike, Suite 1401
Falls Church, VA 22041

SUBJECT: Approval of INTERNATIONAL Single Project Assurance (SPA)

References: (a) 32 CFR 219, Protection of Human Subjects
(b) DoDD 6000.8, Funding and Administration of Clinical Investigation Programs
(c) DoDD 3216.2, Protection of Human Subjects and Adherence to Ethical Standards in DoD-Supported Research
(d) AFI 40-402, Protection of Human Subjects in Research

On behalf of the Air Force Surgeon General, I have approved your INTERNATIONAL SPA application and assigned a SPA number to your institution for the protocol listed below:

SPA Number	USAF Protocol Number and Title
50186	FWR20060063H, "The Utility of Depth Affordance in Multi-Layered Displays for Supporting Change Detection"

Attached is the signed copy of your Assurance of Compliance with the above references. Please maintain these documents with your research records.

We request that you provide this office with a copy of annual status reports and the final report when the project has been completed.

Thank you for your support in this matter. Please do not hesitate to contact me at 703-681-6103, telefax 703- 681-8050, or joe.narrigan@pentagon.af.mil.

JC for Joe Narrigan

JOSEPH J. NARRIGAN, Lt Col, USAF, BSC
Director, Research Oversight/Compliance Division
Office of the Surgeon General

Attachment:
Signed SPA 50186

Print Form

INTERNATIONAL SINGLE PROJECT ASSURANCE

DOD's International Single Project Assurance (ISPA) of Compliance for the Protection of Human Research Subjects

Middlesex University, School of Computing Science, Research Ethics Committee

Middlesex University, School of Computing Science

hereinafter known as the "institution," hereby gives assurance that it will comply with the principles and procedures for protecting human research subjects specified below.

Part 1

Ethical Principles and Institutional Policies Governing Research Involving Human Subjects

I. Applicability

Except when research is exempt from the requirements of the United States Federal Policy for the Protection of Human Research Subjects, or applicability of that policy is waived, this Assurance applies to all research involving human subjects, and all other activities which involve such research even in part, regardless of whether the research is otherwise subject to United States federal regulation, if:

A. Institutional Review Boards (IRBs) [or Institutional Ethics Committees (IECs) as appropriate] operated by the institution provide review and oversight of human subjects research supported by the United States Government, regardless of where the research takes place or by whom it is conducted; or

B. The institution becomes engaged in human subjects research supported by the United States. The institution becomes so engaged whenever

(1) the institution's employees or agents intervene or interact with living individuals for purposes of research supported by the United States;

(2) the institution's employees or agents obtain, release, or access individually identifiable private information for purposes of research supported by the United States; or

(3) the institution receives a direct award to conduct human subjects research supported by the United States, even where all activitiesinvolving human subjects are carried out by a subcontractor or collaborator.

II. Ethical Principles Governing Human Subjects Research

A. This institution is guided by the ethical principles regarding research involving human subjects set forth in the Belmont Report. These ethical principles guide the institution in the conduct of all its human subjects research.

Note: *In Section II above, the institution may choose to cite the Belmont Report, the Declaration of Helsinki, or another appropriate code, declaration, or statement of principles that is consistent with the terms of this Assurance. If choosing another code, select [Other] and attach a brief statement regarding the code to this document.*

B. All U.S. federally-supported human subjects research will comply with the requirements of the Department of Defense (DOD) and any other applicable Federal regulatory agency as well as one or more of the following:

(1) U.S. Department of Health and Human Services (DHHS) regulations at 45 CFR 46 and its Subparts B, C, and D;

(2) The May 1, 1996, International Conference on Harmonization E-6 Guidelines for Good Clinical Practice (ICH-GCP-E6), Sections 1 through 4;

(3) The 1993 Council for International Organizations of Medical Sciences (CIOMS) International Ethical Guidelines for Biomedical Research Involving Human Subjects;

(4) The 1998 Medical Research Council of Canada Tri-Council Policy Statement on Ethical Conduct for Research Involving Humans;

(5) The 2000 Indian Council of Medical Research Ethical Guidelines for Biomedical Research on Human Subjects; or

(6) Other standards for the protection of human subjects recognized by DOD.

C. If the DOD Head determines that the procedures prescribed by the institution afford protections that are at least equivalent to those provided by the U.S. Federal Policy, the DOD Head may approve the substitution of the foreign procedures in lieu of the procedural requirements provided herein, consistent with the requirements of 32 CFR 219.101(h) of the U.S. Federal Policy.

III. Institutional Policies Governing Human Subjects Research

A. This institution acknowledges and accepts its responsibilities for protecting the rights and welfare of all human subjects involved in the research that it sponsors or conducts.

B. This institution encourages and promotes an institutional atmosphere that safeguards the rights and welfare of human subjects.

C. It is the policy of this institution that before human subjects are involved in research that it sponsors or conducts, proper consideration must be given to:

(1) the risks to the subjects,
(2) the anticipated benefits to the subjects and others,
(3) the importance of the knowledge that may reasonably be expected to result, and
(4) the informed consent process to be employed.

D. Whenever appropriate, it is the policy of this institution to consider special safeguards for protecting research subjects who are likely to be vulnerable to coercion or undue influence, such as children, prisoners, pregnant women, mentally disabled persons, or economically or educationally disadvantaged persons. The institution should obtain the concurrence of the supporting U.S. Agency prior to the involvement of pregnant women, prisoners, children, or fetuses in DOD-supported research.

E. This institution encourages and promotes constructive communication among the institutional officials, research administrators, department heads, research investigators, clinical care staff, human subjects, and all other relevant parties as a means of maintaining a high level of awareness regarding the safeguarding of the rights and welfare of the subjects.

Part 2

Human Subjects Protections for Research Funded by the U.S. Air Force

I. Applicability

Part 2 of this Assurance applies only to the following research project that is conducted or sponsored by this institution and supported by the United States Department of Defense (DOD):

Protocol Title: The utility of depth affordance in Multi-Layered Displays for supporting ch

Principal Investigator: Professor William Wong

Address:
School of Computing Science, Ravensfield House,
The Burroughs, Hendon, London NW4 4BT, UK

Tel. +44 (0)208 411 2684 Fax +44 (0)208 411 5216

E-mail w.wong@mdx.ac.uk

II. Institutional Responsibilities

A. This institution recognizes that all human subjects research supported by the U.S. Air Force, including the project referenced above, must be conducted in accordance with the United States Federal Policy for the Protection of Human Research Subjects.

B. The Institutional Review Board (IRB) listed in this Assurance has been designated to be responsible for the initial and continuing review of the project referenced above. The IRB includes at least five persons, including at least one scientist, one non-scientist, and one person not otherwise affiliated with the institution. Every nondiscriminatory effort has been made to include both women and men. The IRB also includes persons who are sensitive to the concerns of the population from which subjects will be recruited.

C. This institution has provided and will continue to provide meeting space for the IRB and sufficient staff to support the IRB's review and record keeping duties.

D. The institution acknowledges that research funded by the United States Department of Defense (DOD) must fully comply with the provisions of Title 10, United States Code, Section 980 (10 USC 980), which states that if an individual cannot give his/her consent (for example, minors), and there is no intent to benefit the individual, he/she cannot be entered into a study funded by the DOD. Furthermore, this institution understands that the DOD is legally bound by United States law to refuse funding of research not in compliance with 10 USC 980, and that if, in the course of institutional oversight of this research, the institution finds the research in violation of 10 USC 980, the institution will immediately notify the DOD of this violation.

E. This institution and the designated IRB (or IEC) have established written procedures for:

(1) verifying whether proposed activities qualify for exemption from, or waiver of, IRB review,
(2) conducting IRB initial and continuing review, approving research, and reporting IRB findings to the investigator and the institution,
(3) determining appropriate continuing review intervals and oversight mechanisms for all approved research,
(4) ensuring that changes in approved research are not initiated without IRB approval, except when necessary to eliminate apparent immediate hazards to the subject, and
(5) ensuring prompt reporting to the IRB, institutional officials, the relevant U.S. Agency Head, and any applicable regulatory body, of any (i) unanticipated problems involving risks to subjects or others in any covered research; (ii) serious or continuing noncompliance with U.S., institutional, or IRB requirements; and (iii) suspension or termination of IRB approval for the DOD-supported research.

F. The institutional Signatory Official, the IRB Administrator and the IRB chairperson will personally complete appropriate relevant human subjects protection and assurance training approved by the DOD approving authority, prior to submitting this Assurance. Members and staff of the IRB will complete relevant training before reviewing human subjects research. Research investigators must complete appropriate institutional training before conducting human subjects research.

G. The institution and the designated IRB have established education and oversight mechanisms (appropriate to the nature and volume of its research) to verify that research investigators, IRB members and staff, and other relevant personnel maintain continuing knowledge of, and comply with, relevant policies and procedures for the protection of human subjects. The institution and the designated IRB will require documentation of such training from research investigators as a condition for conducting DOD-supported human subjects research.

H. Any designation of an IRB(s) not administered by the institution must be documented by a written agreement between the institution and the IRB organization outlining their relationship and including a commitment that the designated IRB will adhere to the requirements of this Assurance.

I. This institution is responsible for ensuring that all institutions and investigators collaborating in its DOD-supported human subjects research operate under an appropriate Assurance of Protection for Human Subjects. All institutions engaged in such research, including subcontractors and subgrantees, must hold their own Assurance.

J. This institution will update information provided under this Assurance **every 36 months**, even if no changes have occurred, in order to maintain an active Assurance. Failure to update this information may result in restriction, suspension, or termination of the institution's Assurance of Protection for Human Subjects.

III. IRB's Responsibilities

A. The project referenced above has been and will be reviewed at convened meetings at which a majority of IRB members are present. A majority vote of those members present at the meeting is required for approval. The research investigators and their family members may not participate in IRB proceedings except to provide information requested by the IRB.

B. The IRB used the following criteria to determine that protections for human research subjects in this project are adequate:

(1) Risks to subjects are minimized.
(2) Risks to subjects are reasonable in relation to anticipated benefits.
(3) Selection of subjects is equitable.
(4) When Appropriate, the data collected will be monitored durning the course of the study to ensure the safety of subjects.
(5) Privacy of subjects and confidentiality of data are protected.

C. The IRB has determined that legally effective written informed consent will be obtained under circumstances that provide sufficient opportunity to consider whether or not to participate and that minimize the possibility of coercion or undue influence. Copies of all informed consent documents for this project will be provided. Informed consent should be written in understandable, nonexculpatory language, and should include the following elements:

(1) Identification as research; purposes, duration, and procedures procedures which are experimental;
(2) Reasonably foreseeable risks or discomforts;
(3) Expected benefits to the subject or others;
(4) Alternative procedures or treatments;
(5) Extent of confidentiality to be maintained;
(6) Whether compensation or medical treatment are available if injury occurs (if more than minimal risk);
(7) Whom to contact for answers to questions about the research, subjects' rights, and research-related injury;
(8) Participation is voluntary; refusal to participate, or discontinuation of participation, will involve no penalty or loss of benefits to which subject is entitled; and
(9) When appropriate, additional elements as determined by the IRB.

D. Where appropriate, the IRB will determine that adequate additional protections are ensured for fetuses, pregnant women, prisoners, and children.

E. The IRB will review, and have the authority to approve, require modification in, or disapprove project changes. The IRB has the authority to suspend or terminate approval of this research activity because of (1) noncompliance with this Assurance document, or the IRB's requirements, or (2) unexpected serious harm to subjects.

F. Continuing reviews by the IRB will be conducted at intervals appropriate to the degree of risk, but not less than once per year. The IRB may be called into an interim review session by the Chairperson at the request of any IRB member or Institutional Official to consider any matter concerned with the rights and welfare of any subject.

G. The IRB will maintain documentation of its activities to include copies of research protocols, minutes of IRB meetings and continuing review records, correspondence with investigators, IRB membership with degrees and affiliations, IRB operating procedures, and statements of new findings provided to subjects. This documentation will be retained for at least three years after the completion of the project and will be accessible for inspection and copying by representatives of the United States Government.

H. The IRB will report promptly to appropriate institutional officials and to the supporting U.S. Government Agency:

(1) Any unanticipated problems or injuries involving risk to subjects or others.
(2) Any serious or continuing noncompliance with this Assurance or with the requirements or determinations of the IRB.
(3) Any changes in this project which are reviewed and approved by the IRB.
(4) Any suspension or termination of IRB approval.

IV. Responsibilities of Principal Investigators

A. Principal investigators accept their responsibility to comply with the stipulations of the IRB and with all applicable provisions of this Assurance.

B. Investigators shall report promptly to the IRB any proposed changes to this project. The changes shall not be initiated without prior IRB approval, except where necessary to eliminate apparent immediate hazards to the subjects.

C. Research investigators shall report promptly to the IRB any unanticipated problems or injuries involving risks to subjects and others.

Part 3

Certification of IRB Approval and Institutional Endorsement

Protocol Title: The utility of depth affordance in Multi-Layered Displays for supporting ch

Principal Investigator: Professor William Wong

Address:
School of Computing Science, Ravensfield House,
The Burroughs, Hendon, London NW4 4BT, UK

Tel. +44 (0)208 411 2684 Fax +44 (0)208 411 5216

E-mail w.wong@mdx.ac.uk

Date of IRB Approval: 5 May 2006

Date of Next IRB Review: 6 mths comp

The officials signing below assure that the project referenced in this Assurance was approved by the IRB on the date indicated above and that the project will be conducted in accordance with all provisions of this Assurance and of the United States Federal Policy for the Protection of Human Research Subjects. A roster listing the current membership of the designated IRB is attached at the end of this Assurance.

I. Authorized Official of the Institution Providing This Assurance

Signature Date 25/05/2006

Professor Colin Tully

Director, Research and Postgraduate Studies, School of Computing Science,

Middlesex University, Ravensfield House,
The Burroughs, Hendon,
London NW4 4BT, UK

Tel. +44 (0)208 411 4616 Fax +44 (0)208 411 5216

E-mail. c.tully@mdx.ac.uk

II. Authorized Official of the Institution with the IRB
(Include only if different from the Institution above)

This institution authorizes the designation of its IRB for review of the project referenced in this Assurance.

Signature: ______________________ Date: OCT 0 2 2006

Hendrick W. Ruck, PhD, SES
Director, Human Effectiveness Directorate
AFRL/HE
2610 Seventh Street
Wright-Patterson AFB, OH 45433-7008
Commercial: 937-255-0215 DSN: 785-0215
Hendrick.Ruck@wpafb.af.mil

III. IRB Chairperson

Signature: ______________________ Date: 21 Sep 06

Jeffrey J. Bidinger, Maj, USAF, MC, FS
Chairperson, Wright Site IRB
AFRL/HEPG
2215 First Street, Bldg 33
Wright-Patterson AFB, OH 45433
937-255-4563
937-255-9687
jeffrey.bidinger@wpafb.af.mil

IV. Responsible Principal Investigator at Institution Providing This Assurance

I have attached copies of IRB approved Informed Consent Documents to be used in this project.

Signature ______________________ Date 26 May 06

Professor William Wong
Head, Interaction Design Centre, School of Computing Science
Middlesex University,
Ravensfield House, The Burroughs,
Hendon, London NW4 4BT, UK
Tel. +44 (0)208 411 2684 Fax +44 (0)208 411 5216
E-mail. w.wong@mdx.ac.uk

DEPARTMENT OF DEFENSE

***International Single Project Assurance (ISPA) Number: DOD A50186**

Middlesex University, School of Computing Science

All parts of this Assurance are in compliance with requirements of Title 32 Code of Federal Regulations Part 219 (32 CFR 219); 10 U.S. Code 980; AFI 40-402; DODD 3216.2; and where applicable 21 CFR 50, 21 CFR 56 and 45 CFR 46 (Subparts B, C, and D) under the authority of the Department of Defense (DOD).

DOD Approving Official

JC for Joe Narrigan   Date: 7 Nov 06

Joe Narrigan, LtCol, USAF, BSC
Director, Research Oversight and Compliance Division

Mailing Address for Communications Regarding this Assurance:

Address: HQ USAF/SGRC
5201 Leesburg Pike, Suite 1401
Falls Church, VA 22041

Telephone: (703)-681-6103
Fax: (703)-681-8050

*Expiration Date: 6 Nov '09

*This assurance expires **three years** from the date of its approval. It must be updated regularly subsequent to a change of the signatory official, the IRB Chair, the IRB membership, or of the policies and procedures to maintain this ISPA file current. A revised and dated IRB membership roster must be submitted if there is a change in the IRB membership. For its uninterrupted continuation, this Assurance must be renegotiated with the Office of the U.S. Air Force Surgeon General via the Chief, HQ USAF/SGRC, prior to its expiration.

Appendix III. Short videos of Experiments 1,2,3 and 4

Appendix IV. Experiment 1 and 2 – Informed Consent Form

School of Computing Science

CONSENT FORM

Using Perceptual Depth to Reduce Change Blindness

I have read the Information Sheet concerning this project and understand what it is about. All my questions have been answered to my satisfaction. I understand that I am free to request further information at any stage.

I know that:

1. My participation in the project is entirely voluntary;

2. I am free to withdraw from the project at any time without any disadvantage;

3. the data *video-tapes and audio-tapes* will be destroyed at the conclusion of the project but any raw data on which the results of the project depend will be retained in secure storage for five years, after which it will be destroyed;

4. This project involves an open-ended questioning technique where the precise nature of the questions that will be asked have not been determined in advance, but will depend on the way in which the interview develops and that in the event that the line of questioning develops in such a way that I feel hesitant or uncomfortable I may decline to answer any particular question(s) and/or may withdraw from the project without any disadvantage of any kind;

5. I will not be exposed to any stress or harm;

6. The results of the project may be published but my anonymity will be preserved.

I agree to take part in this project.

....

(Signature of Participant) (Date)

....

(Signature of Investigator/Witness) (Date)

This project has been reviewed and approved by the Middlesex University School of Computing Science Ethics Committee.

This project is supported by the US Department of Defence

Appendix V. Experiment 1: Repeated Measures ANOVA SPSS Tables

Oneway ANOVA

Descriptives

		N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
						Lower Bound	Upper Bound		
TotalRT	Colour	22	1395.4091	137.95182	29.41143	1334.2447	1456.5735	1203.00	1688.00
	Depth	22	1385.5909	155.00235	33.04661	1316.8667	1454.3151	1078.00	1750.00
	Non-Transient	22	1354.0000	177.56535	37.85706	1275.2719	1432.7281	984.00	1641.00
	Total	66	1378.3333	156.24949	19.23298	1339.9224	1416.7443	984.00	1750.00
TotalCorrect	Colour	22	.8100	.07355	.01568	.7774	.8426	.65	.92
	Depth	22	.9300	.06590	.01405	.9008	.9592	.76	1.00
	Non-Transient	22	.4368	.14597	.03112	.3721	.5015	.14	.71
	Total	66	.7256	.23412	.02882	.6681	.7832	.14	1.00

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
TotalRT	Between Groups	20600.030	2	10300.015	.414	.663
	Within Groups	1566303.636	63	24861.962		
	Total	1586903.667	65			
TotalCorrect	Between Groups	2.911	2	1.455	140.557	.000
	Within Groups	.652	63	.010		
	Total	3.563	65			

Post Hoc Tests

Multiple Comparisons

Tukey HSD

Dependent Variable	(I) TransientType	(J) TransientType	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
TotalRT	Colour	Depth	9.81818	47.54133	.977	-104.2965	123.9329
		Non-Transient	41.40909	47.54133	.660	-72.7056	155.5238
	Depth	Colour	-9.81818	47.54133	.977	-123.9329	104.2965
		Non-Transient	31.59091	47.54133	.785	-82.5238	145.7056
	Non-Transient	Colour	-41.40909	47.54133	.660	-155.5238	72.7056
		Depth	-31.59091	47.54133	.785	-145.7056	82.5238
TotalCorrect	Colour	Depth	-.12000*	.03068	.001	-.1936	-.0464
		Non-Transient	.37318*	.03068	.000	.2995	.4468
	Depth	Colour	.12000*	.03068	.001	.0464	.1936
		Non-Transient	.49318*	.03068	.000	.4195	.5668
	Non-Transient	Colour	-.37318*	.03068	.000	-.4468	-.2995
		Depth	-.49318*	.03068	.000	-.5668	-.4195

*. The mean difference is significant at the 0.05 level.

Homogeneous Subsets

TotalRT

Tukey HSD[a]

TransientType	N	Subset for alpha = 0.05 1
Non-Transient	22	1354.0000
Depth	22	1385.5909
Colour	22	1395.4091
Sig.		.660

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 22.000.

TotalCorrect

Tukey HSD[a]

		Subset for alpha = 0.05		
TransientType	N	1	2	3
Non-Transient	22	.4368		
Colour	22		.8100	
Depth	22			.9300
Sig.		1.000	1.000	1.000

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 22.000.

Accuracy 2x3x3 Repeated Measures

Within-Subjects Factors

Measure:MEASURE_1

transientType	transientDuration	VisualRegion	Dependent Variable
1	1	1	FovealCorrectColor250
		2	ParafovColor250_A
		3	OutpfColor250_A
	2	1	FovealCorrectColor350
		2	ParafovColor350_A
		3	OutpfColor350_A
	3	1	FovealCorrectColor450
		2	ParafovColor450_A
		3	OutpfColor450_A
2	1	1	FovealCorrectDepth250
		2	ParafovDepth250_A
		3	OutpfDepth250_A
	2	1	FovealCorrectDepth350
		2	ParafovDepth350_A
		3	OutpfDepth350_A
	3	1	FovealCorrectDepth450
		2	ParafovDepth450_A
		3	OutpfDepth450_A

Multivariate Tests[c]

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[b]
transientType	Pillai's Trace	.857	107.781[a]	1.000	18.000	.000	.857	107.781	1.000
	Wilks' Lambda	.143	107.781[a]	1.000	18.000	.000	.857	107.781	1.000
	Hotelling's Trace	5.988	107.781[a]	1.000	18.000	.000	.857	107.781	1.000
	Roy's Largest Root	5.988	107.781[a]	1.000	18.000	.000	.857	107.781	1.000
transientDuration	Pillai's Trace	.509	8.801[a]	2.000	17.000	.002	.509	17.603	.939
	Wilks' Lambda	.491	8.801[a]	2.000	17.000	.002	.509	17.603	.939
	Hotelling's Trace	1.035	8.801[a]	2.000	17.000	.002	.509	17.603	.939
	Roy's Largest Root	1.035	8.801[a]	2.000	17.000	.002	.509	17.603	.939
VisualRegion	Pillai's Trace	.854	49.616[a]	2.000	17.000	.000	.854	99.233	1.000
	Wilks' Lambda	.146	49.616[a]	2.000	17.000	.000	.854	99.233	1.000
	Hotelling's Trace	5.837	49.616[a]	2.000	17.000	.000	.854	99.233	1.000
	Roy's Largest Root	5.837	49.616[a]	2.000	17.000	.000	.854	99.233	1.000
transientType * transientDuration	Pillai's Trace	.222	2.420[a]	2.000	17.000	.119	.222	4.839	.420
	Wilks' Lambda	.778	2.420[a]	2.000	17.000	.119	.222	4.839	.420
	Hotelling's Trace	.285	2.420[a]	2.000	17.000	.119	.222	4.839	.420
	Roy's Largest Root	.285	2.420[a]	2.000	17.000	.119	.222	4.839	.420
transientType * VisualRegion	Pillai's Trace	.858	51.512[a]	2.000	17.000	.000	.858	103.024	1.000
	Wilks' Lambda	.142	51.512[a]	2.000	17.000	.000	.858	103.024	1.000
	Hotelling's Trace	6.060	51.512[a]	2.000	17.000	.000	.858	103.024	1.000
	Roy's Largest Root	6.060	51.512[a]	2.000	17.000	.000	.858	103.024	1.000

transientDuration * VisualRegion	Pillai's Trace	.614	5.967[a]	4.000	15.000	.004	.614	23.866	.934
	Wilks' Lambda	.386	5.967[a]	4.000	15.000	.004	.614	23.866	.934
	Hotelling's Trace	1.591	5.967[a]	4.000	15.000	.004	.614	23.866	.934
	Roy's Largest Root	1.591	5.967[a]	4.000	15.000	.004	.614	23.866	.934
transientType * transientDuration * VisualRegion	Pillai's Trace	.366	2.161[a]	4.000	15.000	.123	.366	8.642	.497
	Wilks' Lambda	.634	2.161[a]	4.000	15.000	.123	.366	8.642	.497
	Hotelling's Trace	.576	2.161[a]	4.000	15.000	.123	.366	8.642	.497
	Roy's Largest Root	.576	2.161[a]	4.000	15.000	.123	.366	8.642	.497

a. Exact statistic

b. Computed using alpha = .05

c. Design: Intercept

Within Subjects Design: transientType + transientDuration + VisualRegion + transientType * transientDuration + transientType * VisualRegion + transientDuration * VisualRegion + transientType * transientDuration * VisualRegion

Tests of Within-Subjects Effects

Measure:MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
transientType	Sphericity Assumed	.600	1	.600	107.781	.000	.857	107.781	1.000
	Greenhouse-Geisser	.600	1.000	.600	107.781	.000	.857	107.781	1.000
	Huynh-Feldt	.600	1.000	.600	107.781	.000	.857	107.781	1.000
	Lower-bound	.600	1.000	.600	107.781	.000	.857	107.781	1.000
Error(transientType)	Sphericity Assumed	.100	18	.006					
	Greenhouse-Geisser	.100	18.000	.006					

	Huynh-Feldt	.100	18.000	.006					
	Lower-bound	.100	18.000	.006					
transientDuration	Sphericity Assumed	.109	2	.055	7.376	.002	.291	14.752	.920
	Greenhouse-Geisser	.109	1.874	.058	7.376	.003	.291	13.821	.906
	Huynh-Feldt	.109	2.000	.055	7.376	.002	.291	14.752	.920
	Lower-bound	.109	1.000	.109	7.376	.014	.291	7.376	.729
Error(transientDuration)	Sphericity Assumed	.267	36	.007					
	Greenhouse-Geisser	.267	33.727	.008					
	Huynh-Feldt	.267	36.000	.007					
	Lower-bound	.267	18.000	.015					
VisualRegion	Sphericity Assumed	4.016	2	2.008	100.362	.000	.848	200.723	1.000
	Greenhouse-Geisser	4.016	1.031	3.895	100.362	.000	.848	103.483	1.000
	Huynh-Feldt	4.016	1.037	3.874	100.362	.000	.848	104.040	1.000
	Lower-bound	4.016	1.000	4.016	100.362	.000	.848	100.362	1.000
Error(VisualRegion)	Sphericity Assumed	.720	36	.020					
	Greenhouse-Geisser	.720	18.560	.039					
	Huynh-Feldt	.720	18.660	.039					
	Lower-bound	.720	18.000	.040					
transientType * transientDuration	Sphericity Assumed	.031	2	.015	2.595	.089	.126	5.191	.484
	Greenhouse-Geisser	.031	1.984	.015	2.595	.089	.126	5.150	.482
	Huynh-Feldt	.031	2.000	.015	2.595	.089	.126	5.191	.484
	Lower-bound	.031	1.000	.031	2.595	.125	.126	2.595	.332
Error(transientType*transientDuration)	Sphericity Assumed	.212	36	.006					
	Greenhouse-Geisser	.212	35.717	.006					
	Huynh-Feldt	.212	36.000	.006					

	Lower-bound	.212	18.000	.012					
transientType * VisualRegion	Sphericity Assumed	1.162	2	.581	90.137	.000	.834	180.274	1.000
	Greenhouse-Geisser	1.162	1.119	1.038	90.137	.000	.834	100.880	1.000
	Huynh-Feldt	1.162	1.141	1.018	90.137	.000	.834	102.865	1.000
	Lower-bound	1.162	1.000	1.162	90.137	.000	.834	90.137	1.000
Error(transientType*VisualRegion)	Sphericity Assumed	.232	36	.006					
	Greenhouse-Geisser	.232	20.145	.012					
	Huynh-Feldt	.232	20.542	.011					
	Lower-bound	.232	18.000	.013					
transientDuration * VisualRegion	Sphericity Assumed	.155	4	.039	5.112	.001	.221	20.449	.956
	Greenhouse-Geisser	.155	2.042	.076	5.112	.011	.221	10.437	.796
	Huynh-Feldt	.155	2.305	.067	5.112	.008	.221	11.785	.832
	Lower-bound	.155	1.000	.155	5.112	.036	.221	5.112	.571
Error(transientDuration*VisualRegion)	Sphericity Assumed	.547	72	.008					
	Greenhouse-Geisser	.547	36.748	.015					
	Huynh-Feldt	.547	41.496	.013					
	Lower-bound	.547	18.000	.030					
transientType * transientDuration * VisualRegion	Sphericity Assumed	.068	4	.017	3.114	.020	.147	12.456	.790
	Greenhouse-Geisser	.068	2.185	.031	3.114	.051	.147	6.803	.590
	Huynh-Feldt	.068	2.498	.027	3.114	.044	.147	7.779	.633
	Lower-bound	.068	1.000	.068	3.114	.095	.147	3.114	.386
Error(transientType*transientDuration*VisualRegion)	Sphericity Assumed	.395	72	.005					
	Greenhouse-Geisser	.395	39.325	.010					
	Huynh-Feldt	.395	44.967	.009					
	Lower-bound	.395	18.000	.022					

a

Tests of Within-Subjects Contrasts

Measure:MEASURE_1

Source	transientType	transientDuration	VisualRegion	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
transientType	Linear			.600	1	.600	107.781	.000	.857	107.781	1.000
Error(transientType)	Linear			.100	18	.006					
transientDuration		Linear		.102	1	.102	14.603	.001	.448	14.603	.951
		Quadratic		.007	1	.007	.951	.342	.050	.951	.152
Error(transientDuration)		Linear		.126	18	.007					
		Quadratic		.141	18	.008					
VisualRegion			Linear	3.115	1	3.115	103.921	.000	.852	103.921	1.000
			Quadratic	.901	1	.901	89.740	.000	.833	89.740	1.000
Error(VisualRegion)			Linear	.540	18	.030					
			Quadratic	.181	18	.010					
transientType * transientDuration	Linear	Linear		.017	1	.017	2.686	.119	.130	2.686	.342
		Quadratic		.013	1	.013	2.487	.132	.121	2.487	.321
Error(transientType*transientDuration)	Linear	Linear		.115	18	.006					
		Quadratic		.096	18	.005					
transientType * VisualRegion	Linear		Linear	.881	1	.881	104.651	.000	.853	104.651	1.000
			Quadratic	.281	1	.281	62.845	.000	.777	62.845	1.000

Error(transientType*VisualRegion)	Linear		Linear	.151	18	.008					
			Quadratic	.081	18	.004					
transientDuration * VisualRegion		Linear	Linear	.122	1	.122	12.786	.002	.415	12.786	.922
			Quadratic	.023	1	.023	6.220	.023	.257	6.220	.655
		Quadratic	Linear	.009	1	.009	.739	.401	.039	.739	.129
			Quadratic	.002	1	.002	.305	.588	.017	.305	.082
Error(transientDuration*VisualRegion)		Linear	Linear	.171	18	.010					
			Quadratic	.067	18	.004					
		Quadratic	Linear	.214	18	.012					
			Quadratic	.095	18	.005					
transientType * transientDuration * VisualRegion	Linear	Linear	Linear	.018	1	.018	2.076	.167	.103	2.076	.276
			Quadratic	.002	1	.002	.623	.440	.033	.623	.116
		Quadratic	Linear	.030	1	.030	4.309	.053	.193	4.309	.502
			Quadratic	.018	1	.018	6.835	.018	.275	6.835	.696
Error(transientType*transientDuration*VisualRegion)	Linear	Linear	Linear	.157	18	.009					
			Quadratic	.066	18	.004					
		Quadratic	Linear	.124	18	.007					
			Quadratic	.048	18	.003					

a. Computed using alpha = .05

Tests of Between-Subjects Effects

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
Intercept	289.101	1	289.101	13661.708	.000	.999	13661.708	1.000
Error	.381	18	.021					

a. Computed using alpha = .05

Estimated Marginal Means

1. Grand Mean

Measure:MEASURE_1

Mean	Std. Error	95% Confidence Interval	
		Lower Bound	Upper Bound
.919	.008	.903	.936

2. transientType

Estimates

Measure:MEASURE_1

transient Type	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1	.878	.009	.858	.897

Estimates

Measure:MEASURE_1

transient Type	Mean	Std. Error	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
1	.878	.009	.858	.897
2	.961	.008	.944	.979

Pairwise Comparisons

Measure:MEASURE_1

(I) transient Type	(J) transient Type	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a] Lower Bound	95% Confidence Interval for Difference[a] Upper Bound
1	2	-.084*	.008	.000	-.101	-.067
2	1	.084*	.008	.000	.067	.101

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

a. Adjustment for multiple comparisons: Bonferroni.

Multivariate Tests

	Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[b]
Pillai's trace	.857	107.781[a]	1.000	18.000	.000	.857	107.781	1.000
Wilks' lambda	.143	107.781[a]	1.000	18.000	.000	.857	107.781	1.000
Hotelling's trace	5.988	107.781[a]	1.000	18.000	.000	.857	107.781	1.000

Roy's largest root	5.988	107.781[a]	1.000	18.000	.000	.857	107.781	1.000

Each F tests the multivariate effect of transientType. These tests are based on the linearly independent pairwise comparisons among the estimated marginal means.

a. Exact statistic

b. Computed using alpha = .05

3. transientDuration

Estimates

Measure:MEASURE_1

transient Duration	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1	.902	.012	.877	.927
2	.913	.009	.893	.932
3	.944	.009	.924	.963

Pairwise Comparisons

Measure:MEASURE_1

(I) transient Duration	(J) transient Duration	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a]	
					Lower Bound	Upper Bound
1	2	-.011	.013	1.000	-.045	.022
	3	-.042*	.011	.004	-.071	-.013
2	1	.011	.013	1.000	-.022	.045
	3	-.031*	.010	.022	-.058	-.004

3	1	.042*	.011	.004	.013	.071
	2	.031*	.010	.022	.004	.058

Based on estimated marginal means

a. Adjustment for multiple comparisons: Bonferroni.

*. The mean difference is significant at the .05 level.

Multivariate Tests

	Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[b]
Pillai's trace	.509	8.801[a]	2.000	17.000	.002	.509	17.603	.939
Wilks' lambda	.491	8.801[a]	2.000	17.000	.002	.509	17.603	.939
Hotelling's trace	1.035	8.801[a]	2.000	17.000	.002	.509	17.603	.939
Roy's largest root	1.035	8.801[a]	2.000	17.000	.002	.509	17.603	.939

Each F tests the multivariate effect of transientDuration. These tests are based on the linearly independent pairwise comparisons among the estimated marginal means.

a. Exact statistic

b. Computed using alpha = .05

4. VisualRegion

Estimates

Measure:MEASURE_1

VisualRegion	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1	1.000	.000	1.000	1.000
2	.992	.003	.985	.999

Estimates

Measure:MEASURE_1

VisualRegion	Mean	Std. Error	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
1	1.000	.000	1.000	1.000
2	.992	.003	.985	.999
3	.766	.023	.718	.814

Pairwise Comparisons

Measure:MEASURE_1

(I) VisualRegion	(J) VisualRegion	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a] Lower Bound	95% Confidence Interval for Difference[a] Upper Bound
1	2	.008	.003	.077	.000	.017
	3	.234*	.023	.000	.173	.294
2	1	-.008	.003	.077	-.017	.001
	3	.226*	.023	.000	.166	.286
3	1	-.234*	.023	.000	-.294	-.173
	2	-.226*	.023	.000	-.286	-.166

Based on estimated marginal means

a. Adjustment for multiple comparisons: Bonferroni.

*. The mean difference is significant at the .05 level.

Multivariate Tests

	Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[b]
Pillai's trace	.854	49.616[a]	2.000	17.000	.000	.854	99.233	1.000
Wilks' lambda	.146	49.616[a]	2.000	17.000	.000	.854	99.233	1.000
Hotelling's trace	5.837	49.616[a]	2.000	17.000	.000	.854	99.233	1.000
Roy's largest root	5.837	49.616[a]	2.000	17.000	.000	.854	99.233	1.000

Each F tests the multivariate effect of VisualRegion. These tests are based on the linearly independent pairwise comparisons among the estimated marginal means.

a. Exact statistic

b. Computed using alpha = .05

5. transientType * transientDuration

Measure:MEASURE_1

transient Type	transient Duration	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
1	1	.855	.016	.823	.888
	2	.862	.012	.836	.888
	3	.915	.013	.888	.942
2	1	.948	.014	.918	.977
	2	.964	.010	.942	.985
	3	.973	.008	.956	.989

6. transientType * VisualRegion

Measure:MEASURE_1

transient Type	VisualRegion	Mean	Std. Error	95% Confidence Interval Lower Bound	Upper Bound
1	1	1.000	.000	1.000	1.000
	2	.991	.005	.979	1.002
	3	.642	.029	.582	.702
2	1	1.000	.000	1.000	1.000
	2	.993	.005	.984	1.003
	3	.891	.023	.842	.939

7. transientDuration * VisualRegion

Measure:MEASURE_1

transient Duration	VisualRegion	Mean	Std. Error	95% Confidence Interval Lower Bound	Upper Bound
1	1	1.000	.000	1.000	1.000
	2	.986	.008	.970	1.003
	3	.718	.036	.644	.793
2	1	1.000	.000	1.000	1.000
	2	.990	.006	.978	1.002
	3	.749	.026	.695	.803
3	1	1.000	.000	1.000	1.000

	2	1.000	.000	1.000	1.000
	3	.832	.028	.773	.890

8. transientType * transientDuration * VisualRegion

Measure:MEASURE_1

transient Type	transient Duration	VisualRegion	Mean	Std. Error	95% Confidence Interval	
					Lower Bound	Upper Bound
1	1	1	1.000	.000	1.000	1.000
		2	.978	.015	.945	1.010
		3	.588	.045	.494	.683
	2	1	1.000	.000	1.000	1.000
		2	.994	.006	.982	1.006
		3	.592	.037	.515	.669
	3	1	1.000	.000	1.000	1.000
		2	1.000	.000	1.000	1.000
		3	.745	.039	.663	.827
2	1	1	1.000	.000	1.000	1.000
		2	.995	.005	.984	1.006
		3	.848	.042	.759	.938
	2	1	1.000	.000	1.000	1.000
		2	.985	.010	.963	1.007
		3	.905	.026	.851	.959
	3	1	1.000	.000	1.000	1.000
		2	1.000	.000	1.000	1.000

7. transientDuration * VisualRegion

Measure:MEASURE_1

transient Duration	VisualRe gion	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
1	1	1.000	.000	1.000	1.000
	2	.986	.008	.970	1.003
	3	.718	.036	.644	.793
2	1	1.000	.000	1.000	1.000
	2	.990	.006	.978	1.002
	3	.749	.026	.695	.803
3	1	1.000	.000	1.000	1.000
	2	1.000	.000	1.000	1.000
	3	.918	.024	.868	.968

Post Hoc Comparisons T-Test

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	TotalCorrectColour	.8100	22	.07355	.01568
	TotalCorrectDepth	.9300	22	.06590	.01405
Pair 2	CorrectColourOutPf	.6009	22	.13359	.02848
	CorrectDepthOutPf	.8555	22	.11714	.02497
Pair 3	CorrectColourFovea	.9955	22	.02132	.00455
	CorrectDepthFovea	.9964	22	.01706	.00364
Pair 4	CorrectColourParafovea	.9864	22	.02592	.00553
	CorrectDepthParafovea	.9832	22	.02801	.00597

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	TotalCorrectColour & TotalCorrectDepth	22	.514	.014
Pair 2	CorrectColourOutPf & CorrectDepthOutPf	22	.569	.006
Pair 3	CorrectColourFovea & CorrectDepthFovea	22	-.048	.833
Pair 4	CorrectColourParafovea & CorrectDepthParafovea	22	.528	.011

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
					95% Confidence Interval of the Difference				
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper			
Pair 1	TotalCorrectColour – TotalCorrectDepth	-.12000	.06908	.01473	-.15063	-.08937	-8.148	21	.000
Pair 2	CorrectColourOutPf – CorrectDepthOutPf	-.25455	.11734	.02502	-.30657	-.20252	-10.175	21	.000
Pair 3	CorrectColourFovea – CorrectDepthFovea	-.00091	.02793	.00595	-.01329	.01147	-.153	21	.880
Pair 4	CorrectColourParafovea – CorrectDepthParafovea	.00318	.02626	.00560	-.00846	.01482	.568	21	.576

Response Times 2X3X3 Repeated Measures

Within-Subjects Factors

Measure: MEASURE_1

duration	transients	visualRegions	Dependent Variable
1	1	1	FovealColor250
		2	ParafovColor250
		3	OutpfColor250
	2	1	FovealDepth250
		2	ParafovDepth250
		3	OutpfDepth250
2	1	1	FovealColor350
		2	ParafovColor350
		3	OutpfColor350
	2	1	FovealDepth350

		2	ParafovDepth350
		3	OutpfDepth350
3	1	1	FovealColor450
		2	ParafovColor450
		3	OutpfColor450
	2	1	FovealDepth450
		2	ParafovDepth450
		3	OutpfDepth450

Descriptive Statistics

	Mean	Std. Deviation	N
FovealColor250	1425.8333	478.89871	18
ParafovColor250	1417.4167	199.99068	18
OutpfColor250	1409.7500	202.42204	18
FovealDepth250	1392.6944	250.62816	18
ParafovDepth250	1449.6944	323.83943	18
OutpfDepth250	1351.6667	189.52006	18

FovealColor350	1483.0278	283.81713	18
ParafovColor350	1342.0278	157.94084	18
OutpfColor350	1402.8056	204.11373	18
FovealDepth350	1408.4722	202.16348	18
ParafovDepth350	1355.0278	217.70592	18
OutpfDepth350	1376.2500	115.14764	18
FovealColor450	1422.2778	258.29038	18
ParafovColor450	1382.0278	207.42288	18
OutpfColor450	1443.6389	229.74947	18
FovealDepth450	1431.5278	236.01131	18
ParafovDepth450	1377.0000	185.35364	18
OutpfDepth450	1411.0278	197.84876	18

Multivariate Tests(b)

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
duration	Pillai's Trace	.025	.204(a)	2.000	16.000	.818	.025
	Wilks' Lambda	.975	.204(a)	2.000	16.000	.818	.025
	Hotelling's Trace	.025	.204(a)	2.000	16.000	.818	.025
	Roy's Largest Root	.025	.204(a)	2.000	16.000	.818	.025
transients	Pillai's Trace	.061	1.098(a)	1.000	17.000	.309	.061
	Wilks' Lambda	.939	1.098(a)	1.000	17.000	.309	.061
	Hotelling's Trace	.065	1.098(a)	1.000	17.000	.309	.061
	Roy's Largest Root	.065	1.098(a)	1.000	17.000	.309	.061

visualRegions	Pillai's Trace	.095	.838(a)	2.000	16.000	.451	.095
	Wilks' Lambda	.905	.838(a)	2.000	16.000	.451	.095
	Hotelling's Trace	.105	.838(a)	2.000	16.000	.451	.095
	Roy's Largest Root	.105	.838(a)	2.000	16.000	.451	.095
duration * transients	Pillai's Trace	.006	.052(a)	2.000	16.000	.949	.006
	Wilks' Lambda	.994	.052(a)	2.000	16.000	.949	.006
	Hotelling's Trace	.007	.052(a)	2.000	16.000	.949	.006
	Roy's Largest Root	.007	.052(a)	2.000	16.000	.949	.006
duration * visualRegions	Pillai's Trace	.416	2.489(a)	4.000	14.000	.091	.416
	Wilks' Lambda	.584	2.489(a)	4.000	14.000	.091	.416
	Hotelling's Trace	.711	2.489(a)	4.000	14.000	.091	.416
	Roy's Largest Root	.711	2.489(a)	4.000	14.000	.091	.416
transients * visualRegions	Pillai's Trace	.153	1.446(a)	2.000	16.000	.265	.153
	Wilks' Lambda	.847	1.446(a)	2.000	16.000	.265	.153
	Hotelling's Trace	.181	1.446(a)	2.000	16.000	.265	.153
	Roy's Largest Root	.181	1.446(a)	2.000	16.000	.265	.153
duration * transients * visualRegions	Pillai's Trace	.082	.313(a)	4.000	14.000	.865	.082
	Wilks' Lambda	.918	.313(a)	4.000	14.000	.865	.082
	Hotelling's Trace	.089	.313(a)	4.000	14.000	.865	.082
	Roy's Largest Root	.089	.313(a)	4.000	14.000	.865	.082

a Exact statistic

b Design: Intercept

Within Subjects Design: duration+transients+visualRegions+duration*transients+duration*visualRegions+transients*visualRegions+duration*transients*visualRegions

Mauchly's Test of Sphericity(b)

Measure: MEASURE_1

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon(a)		
					Greenhouse-Geisser	Huynh-Feldt	Lower-bound
duration	.953	.778	2	.678	.955	1.000	.500
transients	1.000	.000	0	.	1.000	1.000	1.000
visualRegions	.515	10.625	2	.005	.673	.710	.500
duration * transients	.823	3.123	2	.210	.849	.934	.500
duration * visualRegions	.270	20.160	9	.018	.626	.743	.250
transients * visualRegions	.599	8.203	2	.017	.714	.761	.500
duration * transients * visualRegions	.338	16.734	9	.054	.650	.778	.250

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

a May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the Tests of Within-Subjects Effects table.

b Design: Intercept

Within Subjects Design: duration+transients+visualRegions+duration*transients+duration*visualRegions+transients*visualRegions+duration*transients*visualRegions

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
duration	Sphericity Assumed	16707.185	2	8353.593	.227	.798	.013
	Greenhouse-Geisser	16707.185	1.909	8750.068	.227	.788	.013
	Huynh-Feldt	16707.185	2.000	8353.593	.227	.798	.013
	Lower-bound	16707.185	1.000	16707.185	.227	.640	.013
Error(duration)	Sphericity Assumed	1248817.287	34	36729.920			
	Greenhouse-Geisser	1248817.287	32.459	38473.182			
	Huynh-Feldt	1248817.287	34.000	36729.920			
	Lower-bound	1248817.287	17.000	73459.840			
transients	Sphericity Assumed	30780.753	1	30780.753	1.098	.309	.061
	Greenhouse-Geisser	30780.753	1.000	30780.753	1.098	.309	.061
	Huynh-Feldt	30780.753	1.000	30780.753	1.098	.309	.061
	Lower-bound	30780.753	1.000	30780.753	1.098	.309	.061
Error(transients)	Sphericity Assumed	476379.414	17	28022.318			

	Greenhouse-Geisser	476379.414	17.000	28022.318			
	Huynh-Feldt	476379.414	17.000	28022.318			
	Lower-bound	476379.414	17.000	28022.318			
visualRegions	Sphericity Assumed	91540.894	2	45770.447	1.420	.256	.077
	Greenhouse-Geisser	91540.894	1.347	67979.607	1.420	.255	.077
	Huynh-Feldt	91540.894	1.421	64433.982	1.420	.256	.077
	Lower-bound	91540.894	1.000	91540.894	1.420	.250	.077
Error(visualRegions)	Sphericity Assumed	1095599.412	34	32223.512			
	Greenhouse-Geisser	1095599.412	22.892	47859.304			
	Huynh-Feldt	1095599.412	24.152	45363.096			
	Lower-bound	1095599.412	17.000	64447.024			
duration * transients	Sphericity Assumed	5351.080	2	2675.540	.047	.954	.003
	Greenhouse-Geisser	5351.080	1.699	3149.911	.047	.933	.003
	Huynh-Feldt	5351.080	1.868	2865.022	.047	.946	.003
	Lower-bound	5351.080	1.000	5351.080	.047	.831	.003
Error(duration*transients)	Sphericity Assumed	1933618.670	34	56871.137			
	Greenhouse-Geisser	1933618.670	28.880	66954.338			
	Huynh-Feldt	1933618.670	31.751	60898.751			
	Lower-bound	1933618.670	17.000	113742.275			
duration * visualRegions	Sphericity Assumed	184759.394	4	46189.848	1.347	.262	.073
	Greenhouse-Geisser	184759.394	2.504	73792.767	1.347	.272	.073
	Huynh-Feldt	184759.394	2.971	62188.542	1.347	.270	.073
	Lower-bound	184759.394	1.000	184759.394	1.347	.262	.073
Error(duration*visualRegions)	Sphericity Assumed	2331705.218	68	34289.783			
	Greenhouse-Geisser	2331705.218	42.564	54781.256			
	Huynh-Feldt	2331705.218	50.506	46166.672			

	Lower-bound	2331705.218	17.000	137159.130			
transients * visualRegions	Sphericity Assumed	44396.048	2	22198.024	.629	.539	.036
	Greenhouse-Geisser	44396.048	1.427	31102.284	.629	.490	.036
	Huynh-Feldt	44396.048	1.521	29179.260	.629	.499	.036
	Lower-bound	44396.048	1.000	44396.048	.629	.439	.036
Error(transients*visualRegions)	Sphericity Assumed	1200138.035	34	35298.178			
	Greenhouse-Geisser	1200138.035	24.266	49457.283			
	Huynh-Feldt	1200138.035	25.865	46399.387			
	Lower-bound	1200138.035	17.000	70596.355			
duration * transients * visualRegions	Sphericity Assumed	37559.035	4	9389.759	.192	.942	.011
	Greenhouse-Geisser	37559.035	2.601	14440.607	.192	.877	.011
	Huynh-Feldt	37559.035	3.112	12067.247	.192	.907	.011
	Lower-bound	37559.035	1.000	37559.035	.192	.667	.011
Error(duration*transients*visualRegions)	Sphericity Assumed	3331518.965	68	48992.926			
	Greenhouse-Geisser	3331518.965	44.216	75346.725			
	Huynh-Feldt	3331518.965	52.912	62963.252			
	Lower-bound	3331518.965	17.000	195971.704			

Tests of Within-Subjects Contrasts

Measure: MEASURE_1

Source	duration	transients	visualRegions	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
duration	Linear			626.963	1	626.963	.020	.888	.001
	Quadratic			16080.222	1	16080.222	.377	.547	.022
Error(duration)	Linear			523778.120	17	30810.478			

	Quadratic			725039.167	17	42649.363			
transients		Linear		30780.753	1	30780.753	1.098	.309	.061
Error(transients)		Linear		476379.414	17	28022.318			
visualRegions			Linear	42686.723	1	42686.723	1.186	.291	.065
			Quadratic	48854.170	1	48854.170	1.717	.208	.092
Error(visualRegions)			Linear	611897.006	17	35993.942			
			Quadratic	483702.406	17	28453.083			
duration * transients	Linear	Linear		1400.463	1	1400.463	.018	.894	.001
	Quadratic	Linear		3950.617	1	3950.617	.108	.747	.006
Error(duration*transients)	Linear	Linear		1308951.454	17	76997.144			
	Quadratic	Linear		624667.216	17	36745.130			
duration * visualRegions	Linear		Linear	7561.752	1	7561.752	.128	.725	.007
			Quadratic	89110.695	1	89110.695	2.025	.173	.106
	Quadratic		Linear	21329.306	1	21329.306	.935	.347	.052
			Quadratic	66757.641	1	66757.641	5.954	.026	.259
Error(duration*visualRegions)	Linear		Linear	1005253.717	17	59132.572			
			Quadratic	747982.628	17	43998.978			
	Quadratic		Linear	387865.308	17	22815.606			
			Quadratic	190603.564	17	11211.974			
transients * visualRegions		Linear	Linear	530.473	1	530.473	.011	.917	.001
			Quadratic	43865.574	1	43865.574	1.936	.182	.102
Error(transients*visualRegions)		Linear	Linear	814955.506	17	47938.559			
			Quadratic	385182.530	17	22657.796			
duration * transients * visualRegions	Linear	Linear	Linear	643.891	1	643.891	.008	.931	.000
			Quadratic	15223.751	1	15223.751	.330	.573	.019
	Quadratic	Linear	Linear	19879.237	1	19879.237	.501	.489	.029
			Quadratic	1812.158	1	1812.158	.065	.802	.004
Error(duration*transients*visualRegions)	Linear	Linear	Linear	1399410.828	17	82318.284			
			Quadratic	783112.489	17	46065.441			
	Quadratic	Linear	Linear	674612.128	17	39683.066			
			Quadratic	474383.519	17	27904.913			

Tests of Between-Subjects Effects

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Intercept	639187951.361	1	639187951.361	1723.205	.000	.990
Error	6305806.750	17	370929.809			

Estimated Marginal Means

1. duration

Estimates

Measure: MEASURE_1

duration	Mean	Std. Error	95% Confidence Interval	
	Lower Bound	Upper Bound	Lower Bound	Upper Bound
1	1407.843	44.878	1313.157	1502.528
2	1394.602	30.359	1330.550	1458.654
3	1411.250	34.336	1338.807	1483.693

Pairwise Comparisons

Measure: MEASURE_1

(I) duration	(J) duration	Mean Difference (I-J)	Std. Error	Sig.(a)	95% Confidence Interval for Difference(a)	
					Lower Bound	Upper Bound
1	2	13.241	28.659	1.000	-62.849	89.330
	3	-3.407	23.886	1.000	-66.826	60.011
2	1	-13.241	28.659	1.000	-89.330	62.849
	3	-16.648	25.468	1.000	-84.267	50.971
3	1	3.407	23.886	1.000	-60.011	66.826
	2	16.648	25.468	1.000	-50.971	84.267

Based on estimated marginal means

a Adjustment for multiple comparisons: Bonferroni.

Multivariate Tests

	Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Pillai's trace	.025	.204(a)	2.000	16.000	.818	.025
Wilks' lambda	.975	.204(a)	2.000	16.000	.818	.025

Hotelling's trace	.025	.204(a)	2.000	16.000	.818	.025
Roy's largest root	.025	.204(a)	2.000	16.000	.818	.025

Each F tests the multivariate effect of duration. These tests are based on the linearly independent pairwise comparisons among the estimated marginal means.

a Exact statistic

2. transients

Estimates

Measure: MEASURE_1

transients	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1	1414.312	33.549	1343.528	1485.095
2	1394.818	36.566	1317.670	1471.966

Pairwise Comparisons

Measure: MEASURE_1

(I) transients	(J) transients	Mean Difference (I-J)	Std. Error	Sig.(a)	95% Confidence Interval for Difference(a)	
					Lower Bound	Upper Bound
1	2	19.494	18.600	.309	-19.748	58.736

2	1	-19.494	18.600	.309	-58.736	19.748

Based on estimated marginal means

a Adjustment for multiple comparisons: Bonferroni.

Multivariate Tests

	Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Pillai's trace	.061	1.098(a)	1.000	17.000	.309	.061
Wilks' lambda	.939	1.098(a)	1.000	17.000	.309	.061
Hotelling's trace	.065	1.098(a)	1.000	17.000	.309	.061
Roy's largest root	.065	1.098(a)	1.000	17.000	.309	.061

Each F tests the multivariate effect of transients. These tests are based on the linearly independent pairwise comparisons among the estimated marginal means.

a Exact statistic

3. duration * transients

Measure: MEASURE_1

duration	transients	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
1	1	1417.667	48.970	1314.350	1520.984
	2	1398.019	49.708	1293.143	1502.894

2	1	1409.287	35.468	1334.456	1484.118
	2	1379.917	37.879	1299.998	1459.835
3	1	1415.981	42.294	1326.749	1505.213
	2	1406.519	38.784	1324.692	1488.345

Appendix VI. Experiment 2: Contingency Tables

Experiment 2: The effect of depth on the detection of unexpected events

Contingency Tables

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
T3_DEC * CONDITION	52	50.0%	52	50.0%	104	100.0%
T4_DEC * CONDITION	52	50.0%	52	50.0%	104	100.0%

T3_DEC * CONDITION

Crosstab

			CONDITION			Total
			SLD	MLD_UEFront	MLD_UEBack	SLD
T3_DEC	YES	Count	5	20	4	29
		Expected Count	6.1	15.1	7.8	29.0
		% within T3_DEC	17.2%	69.0%	13.8%	100.0%
		% within CONDITION	45.5%	74.1%	28.6%	55.8%
	NO	Count	6	7	10	23
		Expected Count	4.9	11.9	6.2	23.0
		% within T3_DEC	26.1%	30.4%	43.5%	100.0%
		% within CONDITION	54.5%	25.9%	71.4%	44.2%
Total		Count	11	27	14	52
		Expected Count	11.0	27.0	14.0	52.0
		% within T3_DEC	21.2%	51.9%	26.9%	100.0%
		% within CONDITION	100.0%	100.0%	100.0%	100.0%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	8.340(a)	2	.015	.013		
Likelihood Ratio	8.581	2	.014	.018		
Fisher's Exact Test	8.254			.013		
Linear-by-Linear Association	2.718(b)	1	.099	.127	.064	.026
N of Valid Cases	52					

a 1 cells (16.7%) have expected count less than 5. The minimum expected count is 4.87.

b The standardized statistic is 1.649.

Symmetric Measures

		Value	Approx. Sig.	Exact Sig.
Nominal by Nominal	Phi	.400	.015	.013
	Cramer's V	.400	.015	.013
N of Valid Cases		52		

a Not assuming the null hypothesis.

b Using the asymptotic standard error assuming the null hypothesis.

T4_DEC * CONDITION

Crosstab

			CONDITION			Total
			SLD	MLD_UEFront	MLD_UEBack	SLD
T4_DEC	YES	Count	8	22	6	36
		Expected Count	7.6	18.7	9.7	36.0
		% within T4_DEC	22.2%	61.1%	16.7%	100.0%
		% within CONDITION	72.7%	81.5%	42.9%	69.2%
	NO	Count	3	5	8	16
		Expected Count	3.4	8.3	4.3	16.0
		% within T4_DEC	18.8%	31.3%	50.0%	100.0%
		% within CONDITION	27.3%	18.5%	57.1%	30.8%
Total		Count	11	27	14	52
		Expected Count	11.0	27.0	14.0	52.0
		% within T4_DEC	21.2%	51.9%	26.9%	100.0%
		% within CONDITION	100.0%	100.0%	100.0%	100.0%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	6.537(a)	2	.038	.043		
Likelihood Ratio	6.306	2	.043	.052		
Fisher's Exact Test	6.150			.048		
Linear-by-Linear Association	4.523(b)	1	.033	.039	.024	.012
N of Valid Cases	52					

a 2 cells (33.3%) have expected count less than 5. The minimum expected count is 3.38.

b The standardized statistic is 2.127.

Symmetric Measures

		Value	Approx. Sig.	Exact Sig.
Nominal by Nominal	Phi	.355	.038	.043
	Cramer's V	.355	.038	.043
N of Valid Cases		52		

a Not assuming the null hypothesis.

b Using the asymptotic standard error assuming the null hypothesis.

Contingency Tables: Eccentricity

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
POS_UE * T3_DEC	52	50.0%	52	50.0%	104	100.0%
POS_UE * T4_DEC	52	50.0%	52	50.0%	104	100.0%

POS_UE * T3_DEC

Crosstab

			T3_DEC		Total
			YES	NO	YES
POS_UE	FAR	Count	9	16	25
		Expected Count	13.9	11.1	25.0
	CENTRE	Count	20	7	27
		Expected Count	15.1	11.9	27.0
Total		Count	29	23	52
		Expected Count	29.0	23.0	52.0

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	7.629(b)	1	.006	.011	.006	
Continuity Correction(a)	6.163	1	.013			
Likelihood Ratio	7.819	1	.005	.011	.006	
Fisher's Exact Test				.011	.006	
Linear-by-Linear Association	7.482(c)	1	.006	.011	.006	.005
N of Valid Cases	52					

a Computed only for a 2x2 table

b 0 cells (.0%) have expected count less than 5. The minimum expected count is 11.06.

c The standardized statistic is -2.735.

Symmetric Measures

		Value	Approx. Sig.	Exact Sig.
Nominal by Nominal	Phi	-.383	.006	.011
	Cramer's V	.383	.006	.011
N of Valid Cases		52		

a Not assuming the null hypothesis.

b Using the asymptotic standard error assuming the null hypothesis.

POS_UE * T4_DEC

Crosstab

			T4_DEC		Total
			YES	NO	YES
POS_UE	FAR	Count	15	10	25
		Expected Count	17.3	7.7	25.0
	CENTRE	Count	21	6	27
		Expected Count	18.7	8.3	27.0
Total		Count	36	16	52
		Expected Count	36.0	16.0	52.0

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	1.926(b)	1	.165	.232	.139	
Continuity Correction(a)	1.182	1	.277			
Likelihood Ratio	1.938	1	.164	.232	.139	
Fisher's Exact Test				.232	.139	
Linear-by-Linear Association	1.889(c)	1	.169	.232	.139	.093
N of Valid Cases	52					

a Computed only for a 2x2 table

b 0 cells (.0%) have expected count less than 5. The minimum expected count is 7.69.

c The standardized statistic is -1.374.

Symmetric Measures

		Value	Approx. Sig.	Exact Sig.
Nominal by Nominal	Phi	-.192	.165	.232
	Cramer's V	.192	.165	.232
N of Valid Cases		52		

a Not assuming the null hypothesis.

b Using the asymptotic standard error assuming the null hypothesis.

Contingency tables: Luminance

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
T3_DEC * COLOR_BALLS	52	50.0%	52	50.0%	104	100.0%
T4_DEC * COLOR_BALLS	52	50.0%	52	50.0%	104	100.0%

T3_DEC * COLOR_BALLS

Crosstab

			COLOR_BALLS		Total
			RED	BLUE	RED
T3_DEC	YES	Count	9	20	29
		Expected Count	13.4	15.6	29.0
		% within T3_DEC	31.0%	69.0%	100.0%
		% within COLOR_BALLS	37.5%	71.4%	55.8%
		% of Total	17.3%	38.5%	55.8%
	NO	Count	15	8	23
		Expected Count	10.6	12.4	23.0
		% within T3_DEC	65.2%	34.8%	100.0%
		% within COLOR_BALLS	62.5%	28.6%	44.2%
		% of Total	28.8%	15.4%	44.2%
Total		Count	24	28	52
		Expected Count	24.0	28.0	52.0
		% within T3_DEC	46.2%	53.8%	100.0%
		% within COLOR_BALLS	100.0%	100.0%	100.0%
		% of Total	46.2%	53.8%	100.0%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	6.031(b)	1	.014	.024	.014	
Continuity Correction(a)	4.734	1	.030			
Likelihood Ratio	6.135	1	.013	.024	.014	
Fisher's Exact Test				.024	.014	
Linear-by-Linear Association	5.915(c)	1	.015	.024	.014	.012
N of Valid Cases	52					

a Computed only for a 2x2 table

b 0 cells (.0%) have expected count less than 5. The minimum expected count is 10.62.

c The standardized statistic is -2.432.

Symmetric Measures

		Value	Approx. Sig.	Exact Sig.
Nominal by Nominal	Phi	-.341	.014	.024
	Cramer's V	.341	.014	.024
N of Valid Cases		52		

a Not assuming the null hypothesis.

b Using the asymptotic standard error assuming the null hypothesis.

T4_DEC * COLOR_BALLS

Crosstab

			COLOR_BALLS		Total
			RED	BLUE	RED
T4_DEC	YES	Count	12	24	36
		Expected Count	16.6	19.4	36.0
		% within T4_DEC	33.3%	66.7%	100.0%
		% within COLOR_BALLS	50.0%	85.7%	69.2%
		% of Total	23.1%	46.2%	69.2%
	NO	Count	12	4	16
		Expected Count	7.4	8.6	16.0
		% within T4_DEC	75.0%	25.0%	100.0%
		% within COLOR_BALLS	50.0%	14.3%	30.8%
		% of Total	23.1%	7.7%	30.8%
Total		Count	24	28	52
		Expected Count	24.0	28.0	52.0
		% within T4_DEC	46.2%	53.8%	100.0%
		% within COLOR_BALLS	100.0%	100.0%	100.0%
		% of Total	46.2%	53.8%	100.0%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	7.738(b)	1	.005	.007	.006	
Continuity Correction(a)	6.152	1	.013			
Likelihood Ratio	7.956	1	.005	.007	.006	
Fisher's Exact Test				.007	.006	
Linear-by-Linear Association	7.589(c)	1	.006	.007	.006	.005
N of Valid Cases	52					

a Computed only for a 2x2 table

b 0 cells (.0%) have expected count less than 5. The minimum expected count is 7.38.

c The standardized statistic is -2.755.

Symmetric Measures

		Value	Approx. Sig.	Exact Sig.
Nominal by Nominal	Phi	-.386	.005	.007
	Cramer's V	.386	.005	.007
N of Valid Cases		52		

a Not assuming the null hypothesis.

b Using the asymptotic standard error assuming the null hypothesis.

Luminance *Detection*Display Type

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
T3_DEC * COLOR_BALLS * CONDITION	52	50.0%	52	50.0%	104	100.0%
T4_DEC * COLOR_BALLS * CONDITION	52	50.0%	52	50.0%	104	100.0%

T3_DEC * COLOR_BALLS * CONDITION

Crosstab

				COLOR_BALLS		Total
CONDITION				RED	BLUE	RED
SLD	T3_DEC	YES	Count	2	3	5
			Expected Count	2.3	2.7	5.0
			% within T3_DEC	40.0%	60.0%	100.0%
			% within COLOR_BALLS	40.0%	50.0%	45.5%
			% of Total	18.2%	27.3%	45.5%
		NO	Count	3	3	6
			Expected Count	2.7	3.3	6.0
			% within T3_DEC	50.0%	50.0%	100.0%
			% within COLOR_BALLS	60.0%	50.0%	54.5%
			% of Total	27.3%	27.3%	54.5%
	Total		Count	5	6	11
			Expected Count	5.0	6.0	11.0
			% within T3_DEC	45.5%	54.5%	100.0%
			% within COLOR_BALLS	100.0%	100.0%	100.0%
			% of Total	45.5%	54.5%	100.0%
MLD_UEFront	T3_DEC	YES	Count	6	14	20
			Expected Count	8.1	11.9	20.0
			% within T3_DEC	30.0%	70.0%	100.0%

			% within COLOR_BALLS	54.5%	87.5%	74.1%
			% of Total	22.2%	51.9%	74.1%
		NO	Count	5	2	7
			Expected Count	2.9	4.1	7.0
			% within T3_DEC	71.4%	28.6%	100.0%
			% within COLOR_BALLS	45.5%	12.5%	25.9%
			% of Total	18.5%	7.4%	25.9%
	Total		Count	11	16	27
			Expected Count	11.0	16.0	27.0
			% within T3_DEC	40.7%	59.3%	100.0%
			% within COLOR_BALLS	100.0%	100.0%	100.0%
			% of Total	40.7%	59.3%	100.0%
MLD_UEBack	T3_DEC	YES	Count	1	3	4
			Expected Count	2.3	1.7	4.0
			% within T3_DEC	25.0%	75.0%	100.0%
			% within COLOR_BALLS	12.5%	50.0%	28.6%
			% of Total	7.1%	21.4%	28.6%
		NO	Count	7	3	10
			Expected Count	5.7	4.3	10.0
			% within T3_DEC	70.0%	30.0%	100.0%
			% within COLOR_BALLS	87.5%	50.0%	71.4%
			% of Total	50.0%	21.4%	71.4%
	Total		Count	8	6	14
			Expected Count	8.0	6.0	14.0
			% within T3_DEC	57.1%	42.9%	100.0%
			% within COLOR_BALLS	100.0%	100.0%	100.0%
			% of Total	57.1%	42.9%	100.0%

Chi-Square Tests

CONDITION		Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
SLD	Pearson Chi-Square	.110(b)	1	.740	1.000	.608	
	Continuity Correction(a)	.000	1	1.000			
	Likelihood Ratio	.110	1	.740	1.000	.608	
	Fisher's Exact Test				1.000	.608	
	Linear-by-Linear Association	.100(c)	1	.752	1.000	.608	.433
	N of Valid Cases	11					
MLD_UEFront	Pearson Chi-Square	3.686(d)	1	.055	.084	.071	
	Continuity Correction(a)	2.170	1	.141			
	Likelihood Ratio	3.688	1	.055	.084	.071	
	Fisher's Exact Test				.084	.071	
	Linear-by-Linear Association	3.550(e)	1	.060	.084	.071	.062
	N of Valid Cases	27					
MLD_UEBack	Pearson Chi-Square	2.363(f)	1	.124	.245	.175	

	Continuity Correction(a)	.882	1	.348			
	Likelihood Ratio	2.405	1	.121	.245	.175	
	Fisher's Exact Test				.245	.175	
	Linear-by-Linear Association	2.194(g)	1	.139	.245	.175	.160
	N of Valid Cases	14					

a Computed only for a 2x2 table
b 4 cells (100.0%) have expected count less than 5. The minimum expected count is 2.27.
c The standardized statistic is -.316.
d 2 cells (50.0%) have expected count less than 5. The minimum expected count is 2.85.
e The standardized statistic is -1.884.
f 3 cells (75.0%) have expected count less than 5. The minimum expected count is 1.71.
g The standardized statistic is -1.481.

Symmetric Measures

CONDITION			Value	Approx. Sig.	Exact Sig.
SLD	Nominal by Nominal	Phi	-.100	.740	1.000
		Cramer's V	.100	.740	1.000
	N of Valid Cases		11		
MLD_UEFront	Nominal by Nominal	Phi	-.369	.055	.084
		Cramer's V	.369	.055	.084
	N of Valid Cases		27		
MLD_UEBack	Nominal by Nominal	Phi	-.411	.124	.245
		Cramer's V	.411	.124	.245
	N of Valid Cases		14		

aNot assuming the null hypothesis.
b Using the asymptotic standard error assuming the null hypothesis.

T4_DEC * COLOR_BALLS * CONDITION

Crosstab

				COLOR_BALLS		Total
CONDITION				RED	BLUE	RED
SLD	T4_DEC	YES	Count	3	5	8
			Expected Count	3.6	4.4	8.0
			% within T4_DEC	37.5%	62.5%	100.0%
			% within COLOR_BALLS	60.0%	83.3%	72.7%
			% of Total	27.3%	45.5%	72.7%
		NO	Count	2	1	3
			Expected Count	1.4	1.6	3.0
			% within T4_DEC	66.7%	33.3%	100.0%
			% within COLOR_BALLS	40.0%	16.7%	27.3%
			% of Total	18.2%	9.1%	27.3%
	Total		Count	5	6	11
			Expected Count	5.0	6.0	11.0

			% within T4_DEC	45.5%	54.5%	100.0%
			% within COLOR_BALLS	100.0%	100.0%	100.0%
			% of Total	45.5%	54.5%	100.0%
MLD_UEFront	T4_DEC	YES	Count	6	16	22
			Expected Count	9.0	13.0	22.0
			% within T4_DEC	27.3%	72.7%	100.0%
			% within COLOR_BALLS	54.5%	100.0%	81.5%
			% of Total	22.2%	59.3%	81.5%
		NO	Count	5	0	5
			Expected Count	2.0	3.0	5.0
			% within T4_DEC	100.0%	.0%	100.0%
			% within COLOR_BALLS	45.5%	.0%	18.5%
			% of Total	18.5%	.0%	18.5%
	Total		Count	11	16	27
			Expected Count	11.0	16.0	27.0
			% within T4_DEC	40.7%	59.3%	100.0%
			% within COLOR_BALLS	100.0%	100.0%	100.0%
			% of Total	40.7%	59.3%	100.0%
MLD_UEBack	T4_DEC	YES	Count	3	3	6
			Expected Count	3.4	2.6	6.0
			% within T4_DEC	50.0%	50.0%	100.0%
			% within COLOR_BALLS	37.5%	50.0%	42.9%
			% of Total	21.4%	21.4%	42.9%
		NO	Count	5	3	8
			Expected Count	4.6	3.4	8.0
			% within T4_DEC	62.5%	37.5%	100.0%
			% within COLOR_BALLS	62.5%	50.0%	57.1%
			% of Total	35.7%	21.4%	57.1%
	Total		Count	8	6	14
			Expected Count	8.0	6.0	14.0
			% within T4_DEC	57.1%	42.9%	100.0%
			% within COLOR_BALLS	100.0%	100.0%	100.0%
			% of Total	57.1%	42.9%	100.0%

Chi-Square Tests

CONDITION		Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
SLD	Pearson Chi-Square	.749(b)	1	.387	.545	.424	
	Continuity Correction(a)	.034	1	.853			
	Likelihood Ratio	.754	1	.385	.545	.424	
	Fisher's Exact Test				.545	.424	
	Linear-by-Linear Association	.681(c)	1	.409	.545	.424	.364
	N of Valid Cases	11					
MLD_UEFront	Pearson Chi-Square	8.926(d)	1	.003	.006	.006	
	Continuity Correction(a)	6.167	1	.013			

	Likelihood Ratio	10.717	1	.001	.006	.006	
	Fisher's Exact Test				.006	.006	
	Linear-by-Linear Association	8.595(e)	1	.003	.006	.006	.006
	N of Valid Cases	27					
MLD_UEBack	Pearson Chi-Square	.219(f)	1	.640	1.000	.529	
	Continuity Correction(a)	.000	1	1.000			
	Likelihood Ratio	.219	1	.640	1.000	.529	
	Fisher's Exact Test				1.000	.529	
	Linear-by-Linear Association	.203(g)	1	.652	1.000	.529	.373
	N of Valid Cases	14					

a Computed only for a 2x2 table
b 4 cells (100.0%) have expected count less than 5. The minimum expected count is 1.36.
c The standardized statistic is -.825.
d 2 cells (50.0%) have expected count less than 5. The minimum expected count is 2.04.
e The standardized statistic is -2.932.
f 4 cells (100.0%) have expected count less than 5. The minimum expected count is 2.57.
g The standardized statistic is -.451.

Symmetric Measures

CONDITION			Value	Approx. Sig.	Exact Sig.
SLD	Nominal by Nominal	Phi	-.261	.387	.545
		Cramer's V	.261	.387	.545
	N of Valid Cases		11		
MLD_UEFront	Nominal by Nominal	Phi	-.575	.003	.006
		Cramer's V	.575	.003	.006
	N of Valid Cases		27		
MLD_UEBack	Nominal by Nominal	Phi	-.125	.640	1.000
		Cramer's V	.125	.640	1.000
	N of Valid Cases		14		

a Not assuming the null hypothesis.

b Using the asymptotic standard error assuming the null hypothesis.

Appendix VII. Experiment 3 and 4 – Informed Consent Form

School of Computing Science

Participant information –

Evaluating the Depth of the MLD for Comparisons Tasks

The experiment you are about to participate in is investigating user performance in comparison and detection tasks using a Multi Layered Display (MLD). This display has two LCD screens, one in front of the other, and is separated by 14 mm. This allows to present information in different depth layers.

You will be asked to find the difference between two images. One difference will be always present. The images in Experiment 1 will be simple shapes. The images in Experiment 2 will include shapes, cartoons, photographs and Excel graphs. These images will be presented side by side, or, one in front of the other. You will be asked to click on the difference. If you are unable to see the difference, please click the right button. Images will automatically disappear after 20 seconds if no response has been made. Additionally we will ask you to fill out a questionnaire.

Duration of the experiment will be between 45 minutes to one hour.

Looking at computer screens for a long time can produce fatigue and/or eye strain. If you feel any discomfort, please take a break and inform the experimenter. If you feel that you cannot continue, please advice the experimenter immediately and the simulation will cease.

All information and data collected in this study is kept private and confidential to the experimenter. Individual participants will not be identifiable in results or publications. Your participation in this study is voluntary and you may withdraw at any time, including the withdrawal of any information you have provided. However, by signing the consent form attached, it is understood that you have consented to participate in this experiment and in the publication of the results, with the understanding that confidentiality will be preserved.

If you have questions related to the study or the results obtained from your participation, please feel free to contact Gabriela Mancero at g.mancero@mdx.ac.uk or at 020 8411 4981

Please take this sheet with you when you leave.

General data

Please circle the appropriate answer or fill in the spaces provided:

1. Sex: M / F
2. Age (years): _______
3. Eyesight problems / defective vision: yes / no

   If yes, please describe: ____

   Is it corrected (do you wear glasses or contact lenses, etc.)? yes / no

4. How many hours per day do you normally use a computer (watch a computer screen):

   ____

5. How many hours per week do you play computer/video games:

   ____

6. Have you already participated in a similar task? If yes, please describe:

   ____

   ____

Thank you for your participation and cooperation.

Gabriela Mancero

PhD Research Student

Appendix VIII. Experiment 3: Repeated Measures ANOVA SPSS Tables

Experiment 3: Comparing Simple Stimuli in the MLD

2 x 6 Repeated Measures ANOVA Response Times

Within-Subjects Factors

Measure: MEASURE_1

DisplayType	ManipulationType	Dependent Variable
1	1	SldHue
	2	SLDShade
	3	SLDGrow
	4	SLDShrink
	5	SLDHoleAdd
	6	SLDHoleFill
2	1	MLDHue
	2	MLDShade
	3	MLDGrow
	4	MLDShrink
	5	MLDHoleAdd
	6	MLDHoleFill

Descriptive Statistics

	Mean	Std. Deviation	N
SldHue	5697.1389	2296.18177	18
SLDShade	6182.0000	2906.43447	18
SLDGrow	6603.0278	2237.81174	18
SLDShrink	5435.9444	2431.96051	18
SLDHoleAdd	1937.4444	649.91008	18
SLDHoleFill	1819.3333	448.31656	18
MLDHue	6190.8611	3907.31505	18
MLDShade	7470.3889	3674.11767	18
MLDGrow	2904.7222	1720.92857	18
MLDShrink	1800.7778	594.42454	18
MLDHoleAdd	2303.1389	1034.97723	18
MLDHoleFill	2189.5278	1197.68116	18

Multivariate Tests(c)

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power(a)
DisplayType	Pillai's Trace	.455	14.185(b)	1.000	17.000	.002	.455	14.185	.944
	Wilks' Lambda	.545	14.185(b)	1.000	17.000	.002	.455	14.185	.944
	Hotelling's Trace	.834	14.185(b)	1.000	17.000	.002	.455	14.185	.944
	Roy's Largest Root	.834	14.185(b)	1.000	17.000	.002	.455	14.185	.944
ManipulationType	Pillai's Trace	.914	27.636(b)	5.000	13.000	.000	.914	138.182	1.000

	Wilks' Lambda	.086	27.636(b)	5.000	13.000	.000	.914	138.182	1.000
	Hotelling's Trace	10.629	27.636(b)	5.000	13.000	.000	.914	138.182	1.000
	Roy's Largest Root	10.629	27.636(b)	5.000	13.000	.000	.914	138.182	1.000
DisplayType * ManipulationType	Pillai's Trace	.824	12.167(b)	5.000	13.000	.000	.824	60.837	1.000
	Wilks' Lambda	.176	12.167(b)	5.000	13.000	.000	.824	60.837	1.000
	Hotelling's Trace	4.680	12.167(b)	5.000	13.000	.000	.824	60.837	1.000
	Roy's Largest Root	4.680	12.167(b)	5.000	13.000	.000	.824	60.837	1.000

a Computed using alpha = .05

b Exact statistic

c Design: Intercept

Within Subjects Design: DisplayType+ManipulationType+DisplayType*ManipulationType

Mauchly's Test of Sphericity(b)

Measure: MEASURE_1

	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon(a)		
Within Subjects Effect	Greenhouse-Geisser	Huynh-Feldt	Lower-bound	Greenhouse-Geisser	Huynh-Feldt	Lower-bound	Greenhouse-Geisser
DisplayType	1.000	.000	0	.	1.000	1.000	1.000
ManipulationType	.011	67.770	14	.000	.444	.513	.200
DisplayType * ManipulationType	.021	58.431	14	.000	.619	.773	.200

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

a May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the Tests of Within-Subjects Effects table.

b Design: Intercept

Within Subjects Design: DisplayType+ManipulationType+DisplayType*ManipulationType

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power(a)
DisplayType	Sphericity Assumed	34783159.084	1	34783159.084	14.185	.002	.455	14.185	.944
	Greenhouse-Geisser	34783159.084	1.000	34783159.084	14.185	.002	.455	14.185	.944
	Huynh-Feldt	34783159.084	1.000	34783159.084	14.185	.002	.455	14.185	.944
	Lower-bound	34783159.084	1.000	34783159.084	14.185	.002	.455	14.185	.944
Error(DisplayType)	Sphericity Assumed	41686586.061	17	2452152.121					
	Greenhouse-Geisser	41686586.061	17.000	2452152.121					
	Huynh-Feldt	41686586.061	17.000	2452152.121					
	Lower-bound	41686586.061	17.000	2452152.121					
ManipulationType	Sphericity Assumed	710225097.381	5	142045019.476	33.870	.000	.666	169.351	1.000
	Greenhouse-Geisser	710225097.381	2.219	320061120.747	33.870	.000	.666	75.159	1.000
	Huynh-Feldt	710225097.381	2.567	276676737.868	33.870	.000	.666	86.944	1.000
	Lower-bound	710225097.381	1.000	710225097.381	33.870	.000	.666	33.870	1.000
Error(ManipulationType)	Sphericity Assumed	356473913.682	85	4193810.749					
	Greenhouse-Geisser	356473913.682	37.724	9449650.354					
	Huynh-Feldt	356473913.682	43.639	8168747.356					
	Lower-bound	356473913.68	17.000	20969053.746					

		2							
DisplayType * ManipulationType	Sphericity Assumed	226814302.242	5	45362860.448	11.778	.000	.409	58.888	1.000
	Greenhouse-Geisser	226814302.242	3.096	73263968.593	11.778	.000	.409	36.462	.999
	Huynh-Feldt	226814302.242	3.864	58699705.556	11.778	.000	.409	45.508	1.000
	Lower-bound	226814302.242	1.000	226814302.242	11.778	.003	.409	11.778	.898
Error(DisplayType*ManipulationType)	Sphericity Assumed	327388886.987	85	3851633.965					
	Greenhouse-Geisser	327388886.987	52.629	6220639.241					
	Huynh-Feldt	327388886.987	65.688	4984028.286					
	Lower-bound	327388886.987	17.000	19258169.823					

a Computed using alpha = .05

Tests of Within-Subjects Contrasts

Measure: MEASURE_1

Source	DisplayType	ManipulationType	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power(a)
DisplayType	Linear		34783159.084	1	34783159.084	14.185	.002	.455	14.185	.944
Error(DisplayType)	Linear		41686586.061	17	2452152.121					
ManipulationType		Linear	628239731.302	1	628239731.302	67.463	.000	.799	67.463	1.000
		Quadratic	3108747.938	1	3108747.938	.882	.361	.049	.882	.144
		Cubic	63264999.667	1	63264999.667	44.685	.000	.724	44.685	1.000
		Order 4	5924186.683	1	5924186.683	1.359	.260	.074	1.359	.196

		Order 5	9687431.791	1	9687431.791	4.108	.059	.195	4.108	.481
Error(ManipulationType)		Linear	158309090.246	17	9312299.426					
		Quadratic	59920586.857	17	3524740.403					
		Cubic	24068537.251	17	1415796.309					
		Order 4	74083181.898	17	4357834.229					
		Order 5	40092517.430	17	2358383.378					
DisplayType * ManipulationType	Linear	Linear	1419372.001	1	1419372.001	.416	.528	.024	.416	.093
		Quadratic	109710095.278	1	109710095.278	30.412	.000	.641	30.412	.999
		Cubic	1561659.756	1	1561659.756	.458	.508	.026	.458	.098
		Order 4	113186458.953	1	113186458.953	24.368	.000	.589	24.368	.996
		Order 5	936716.254	1	936716.254	.224	.642	.013	.224	.073
Error(DisplayType*ManipulationType)	Linear	Linear	58047046.819	17	3414532.166					
		Quadratic	61327454.147	17	3607497.303					
		Cubic	57938856.796	17	3408168.047					
		Order 4	78962385.413	17	4644846.201					
		Order 5	71113143.813	17	4183126.107					

a Computed using alpha = .05

Tests of Between-Subjects Effects

Measure: MEASURE_1
Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power(a)
Intercept	3830574056.973	1	3830574056.973	223.060	.000	.929	223.060	1.000

Error	291938379.339	17	17172845.843					

a Computed using alpha = .05

Estimated Marginal Means

1. Grand Mean

Measure: MEASURE_1

Mean	Std. Error	95% Confidence Interval	
		Lower Bound	Upper Bound
4211.192	281.964	3616.299	4806.085

2. DisplayType

Estimates

Measure: MEASURE_1

DisplayType	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1	4612.481	243.313	4099.135	5125.828
2	3809.903	350.017	3071.432	4548.373

Pairwise Comparisons

Measure: MEASURE_1

(I) DisplayType	(J) DisplayType	Mean Difference (I-J)	Std. Error	Sig.(a)	95% Confidence Interval for Difference(a)	
					Lower Bound	Upper Bound

1	2	802.579(*)	213.097	.002	352.984	1252.174
2	1	-802.579(*)	213.097	.002	-1252.174	-352.984

Based on estimated marginal means

* The mean difference is significant at the .05 level.

a Adjustment for multiple comparisons: Bonferroni.

Multivariate Tests

	Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power(a)
Pillai's trace	.455	14.185(b)	1.000	17.000	.002	.455	14.185	.944
Wilks' lambda	.545	14.185(b)	1.000	17.000	.002	.455	14.185	.944
Hotelling's trace	.834	14.185(b)	1.000	17.000	.002	.455	14.185	.944
Roy's largest root	.834	14.185(b)	1.000	17.000	.002	.455	14.185	.944

Each F tests the multivariate effect of DisplayType. These tests are based on the linearly independent pairwise comparisons among the estimated marginal means.

a Computed using alpha = .05

b Exact statistic

3. ManipulationType

Estimates

Measure: MEASURE_1

ManipulationType	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1	5944.000	640.888	4591.846	7296.154
2	6826.194	630.433	5496.097	8156.292

3	4753.875	330.805	4055.937	5451.813
4	3618.361	304.711	2975.476	4261.246
5	2120.292	146.492	1811.220	2429.363
6	2004.431	166.053	1654.089	2354.772

Pairwise Comparisons

Measure: MEASURE_1

(I) ManipulationType	(J) ManipulationType	Mean Difference (I-J)	Std. Error	Sig.(a)	95% Confidence Interval for Difference(a)	
					Lower Bound	Upper Bound
1	2	-882.194	475.601	1.000	-2504.093	739.704
	3	1190.125	531.132	.580	-621.142	3001.392
	4	2325.639(*)	630.073	.027	176.962	4474.316
	5	3823.708(*)	609.279	.000	1745.943	5901.473
	6	3939.569(*)	562.383	.000	2021.728	5857.410
2	1	882.194	475.601	1.000	-739.704	2504.093
	3	2072.319	655.281	.085	-162.324	4306.963
	4	3207.833(*)	620.351	.001	1092.309	5323.357
	5	4705.903(*)	576.064	.000	2741.405	6670.401
	6	4821.764(*)	545.562	.000	2961.284	6682.244
3	1	-1190.125	531.132	.580	-3001.392	621.142
	2	-2072.319	655.281	.085	-4306.963	162.324
	4	1135.514(*)	280.895	.013	177.606	2093.422
	5	2633.583(*)	329.328	.000	1510.509	3756.658

	6	2749.444(*)	322.387	.000	1650.040	3848.849
4	1	-2325.639(*)	630.073	.027	-4474.316	-176.962
	2	-3207.833(*)	620.351	.001	-5323.357	-1092.309
	3	-1135.514(*)	280.895	.013	-2093.422	-177.606
	5	1498.069(*)	280.672	.001	540.919	2455.220
	6	1613.931(*)	277.397	.000	667.950	2559.911
5	1	-3823.708(*)	609.279	.000	-5901.473	-1745.943
	2	-4705.903(*)	576.064	.000	-6670.401	-2741.405
	3	-2633.583(*)	329.328	.000	-3756.658	-1510.509
	4	-1498.069(*)	280.672	.001	-2455.220	-540.919
	6	115.861	106.409	1.000	-247.014	478.737
6	1	-3939.569(*)	562.383	.000	-5857.410	-2021.728
	2	-4821.764(*)	545.562	.000	-6682.244	-2961.284
	3	-2749.444(*)	322.387	.000	-3848.849	-1650.040
	4	-1613.931(*)	277.397	.000	-2559.911	-667.950
	5	-115.861	106.409	1.000	-478.737	247.014

Based on estimated marginal means

* The mean difference is significant at the .05 level.

a Adjustment for multiple comparisons: Bonferroni.

Multivariate Tests

	Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power(a)
Pillai's trace	.914	27.636(b)	5.000	13.000	.000	.914	138.182	1.000
Wilks' lambda	.086	27.636(b)	5.000	13.000	.000	.914	138.182	1.000

Hotelling's trace	10.629	27.636(b)	5.000	13.000	.000	.914	138.182	1.000
Roy's largest root	10.629	27.636(b)	5.000	13.000	.000	.914	138.182	1.000

Each F tests the multivariate effect of ManipulationType. These tests are based on the linearly independent pairwise comparisons among the estimated marginal means.

a Computed using alpha = .05

b Exact statistic

4. DisplayType * ManipulationType

Measure: MEASURE_1

DisplayType	ManipulationType	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
1	1	5697.139	541.215	4555.275	6839.003
	2	6182.000	685.053	4736.664	7627.336
	3	6603.028	527.457	5490.190	7715.865
	4	5435.944	573.219	4226.559	6645.330
	5	1937.444	153.185	1614.252	2260.637
	6	1819.333	105.669	1596.391	2042.276
2	1	6190.861	920.963	4247.799	8133.923
	2	7470.389	865.998	5643.293	9297.485
	3	2904.722	405.627	2048.925	3760.520
	4	1800.778	140.107	1505.177	2096.378
	5	2303.139	243.946	1788.457	2817.821
	6	2189.528	282.296	1593.935	2785.121

2 x 6 Repeated Measures ANOVA - Accuracy

Within-Subjects Factors

Measure: MEASURE_1

DisplayType	ManipulationType	Dependent Variable
1	1	SLDCorrectHUES
	2	SLDCorrectSHADES
	3	SLDCorrectGROW
	4	SLDCorrectSHRINK
	5	SLDCorrectHOLEADD
	6	SLDCorrectHOLEFILL
2	1	MLDCorrectHUE
	2	MLDCorrectSHADE
	3	MLDCorrectGROW
	4	MLDCorrectSHRINK

5	MLDCorrectHOLEADD
6	MLDCorrectHOLEFILL

Descriptive Statistics

	Mean	Std. Deviation	N
SLDCorrectHUES	.9022	.14576	23
SLDCorrectSHADES	.7826	.15639	23
SLDCorrectGROW	.7983	.21854	23
SLDCorrectSHRINK	.8991	.18579	23
SLDCorrectHOLEADD	.9852	.04898	23
SLDCorrectHOLEFILL	.9926	.03545	23
MLDCorrectHUE	.5435	.25730	23
MLDCorrectSHADE	.3261	.17573	23
MLDCorrectGROW	.9274	.20062	23
MLDCorrectSHRINK	.9135	.17994	23
MLDCorrectHOLEADD	.8983	.19225	23
MLDCorrectHOLEFILL	.9413	.09607	23

Multivariate Tests(c)

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power(a)
DisplayType	Pillai's Trace	.430	16.574(b)	1.000	22.000	.001	.430	16.574	.973
	Wilks' Lambda	.570	16.574(b)	1.000	22.000	.001	.430	16.574	.973
	Hotelling's Trace	.753	16.574(b)	1.000	22.000	.001	.430	16.574	.973
	Roy's Largest Root	.753	16.574(b)	1.000	22.000	.001	.430	16.574	.973
ManipulationType	Pillai's Trace	.930	47.732(b)	5.000	18.000	.000	.930	238.660	1.000
	Wilks' Lambda	.070	47.732(b)	5.000	18.000	.000	.930	238.660	1.000
	Hotelling's Trace	13.259	47.732(b)	5.000	18.000	.000	.930	238.660	1.000
	Roy's Largest Root	13.259	47.732(b)	5.000	18.000	.000	.930	238.660	1.000
DisplayType * ManipulationType	Pillai's Trace	.876	25.344(b)	5.000	18.000	.000	.876	126.722	1.000
	Wilks' Lambda	.124	25.344(b)	5.000	18.000	.000	.876	126.722	1.000
	Hotelling's Trace	7.040	25.344(b)	5.000	18.000	.000	.876	126.722	1.000
	Roy's Largest Root	7.040	25.344(b)	5.000	18.000	.000	.876	126.722	1.000

a Computed using alpha = .05

b Exact statistic

c Design: Intercept
Within Subjects Design: DisplayType+ManipulationType+DisplayType*ManipulationType

Mauchly's Test of Sphericity(b)

Measure: MEASURE_1

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon(a)		
					Greenhouse-Geisser	Huynh-Feldt	Lower-bound
DisplayType	1.000	.000	0	.	1.000	1.000	1.000
ManipulationType	.362	20.432	14	.119	.779	.967	.200
DisplayType * ManipulationType	.256	27.367	14	.018	.725	.886	.200

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

a May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the Tests of Within-Subjects Effects table.

b Design: Intercept

Within Subjects Design: DisplayType+ManipulationType+DisplayType*ManipulationType

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power(a)
DisplayType	Sphericity Assumed	1.258	1	1.258	16.574	.001	.430	16.574	.973
	Greenhouse-Geisser	1.258	1.000	1.258	16.574	.001	.430	16.574	.973
	Huynh-Feldt	1.258	1.000	1.258	16.574	.001	.430	16.574	.973
	Lower-bound	1.258	1.000	1.258	16.574	.001	.430	16.574	.973
Error(DisplayType)	Sphericity Assumed	1.669	22	.076					
	Greenhouse-Geisser	1.669	22.000	.076					
	Huynh-Feldt	1.669	22.000	.076					
	Lower-bound	1.669	22.000	.076					
ManipulationType	Sphericity Assumed	5.773	5	1.155	52.712	.000	.706	263.562	1.000
	Greenhouse-Geisser	5.773	3.895	1.482	52.712	.000	.706	205.293	1.000
	Huynh-Feldt	5.773	4.837	1.194	52.712	.000	.706	254.968	1.000
	Lower-bound	5.773	1.000	5.773	52.712	.000	.706	52.712	1.000
Error(ManipulationType)	Sphericity Assumed	2.410	110	.022					
	Greenhouse-Geisser	2.410	85.681	.028					
	Huynh-Feldt	2.410	106.413	.023					
	Lower-bound	2.410	22.000	.110					
DisplayType * ManipulationType	Sphericity Assumed	2.930	5	.586	25.367	.000	.536	126.834	1.000
	Greenhouse-Geisser	2.930	3.626	.808	25.367	.000	.536	91.977	1.000
	Huynh-Feldt	2.930	4.430	.661	25.367	.000	.536	112.371	1.000
	Lower-bound	2.930	1.000	2.930	25.367	.000	.536	25.367	.998
Error(DisplayType*Manipulati	Sphericity Assumed	2.541	110	.023					

onType)	Greenhouse-Geisser	2.541	79.769	.032					
	Huynh-Feldt	2.541	97.457	.026					
	Lower-bound	2.541	22.000	.116					

a Computed using alpha = .05

Tests of Within-Subjects Contrasts

Measure: MEASURE_1

Source	DisplayType	ManipulationType	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power(a)
DisplayType	Linear		1.258	1	1.258	16.574	.001	.430	16.574	.973
Error(DisplayType)	Linear		1.669	22	.076					
ManipulationType		Linear	3.869	1	3.869	177.030	.000	.889	177.030	1.000
		Quadratic	.008	1	.008	.419	.524	.019	.419	.095
		Cubic	.708	1	.708	35.985	.000	.621	35.985	1.000
		Order 4	.899	1	.899	49.742	.000	.693	49.742	1.000
		Order 5	.289	1	.289	9.655	.005	.305	9.655	.844
Error(ManipulationType)		Linear	.481	22	.022					
		Quadratic	.440	22	.020					
		Cubic	.433	22	.020					
		Order 4	.398	22	.018					
		Order 5	.658	22	.030					
DisplayType * ManipulationType	Linear	Linear	1.052	1	1.052	50.000	.000	.694	50.000	1.000
		Quadratic	.593	1	.593	46.226	.000	.678	46.226	1.000
		Cubic	.022	1	.022	.757	.394	.033	.757	.132
		Order 4	.933	1	.933	55.907	.000	.718	55.907	1.000
		Order 5	.330	1	.330	9.297	.006	.297	9.297	.830

Error(DisplayType*ManipulationType)	Linear	Linear	.463	22	.021					
		Quadratic	.282	22	.013					
		Cubic	.649	22	.029					
		Order 4	.367	22	.017					
		Order 5	.780	22	.035					

a Computed using alpha = .05

Tests of Between-Subjects Effects

Measure: MEASURE_1
Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power(a)
Intercept	188.232	1	188.232	4037.086	.000	.995	4037.086	1.000
Error	1.026	22	.047					

a Computed using alpha = .05

Estimated Marginal Means

1. Grand Mean

Measure: MEASURE_1

Mean	Std. Error	95% Confidence Interval	
		Lower Bound	Upper Bound
.826	.013	.799	.853

2. DisplayType

Estimates

Measure: MEASURE_1

DisplayType	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1	.893	.013	.867	.920
2	.758	.027	.703	.814

Pairwise Comparisons

Measure: MEASURE_1

(I) DisplayType	(J) DisplayType	Mean Difference (I-J)	Std. Error	Sig.(a)	95% Confidence Interval for Difference(a)	
					Lower Bound	Upper Bound
1	2			.001	.066	.204
2	1	-.135(*)	.033	.001	-.204	-.066

Based on estimated marginal means

* The mean difference is significant at the .05 level.

a Adjustment for multiple comparisons: Bonferroni.

Multivariate Tests

	Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power(a)
Pillai's trace	.430	16.574(b)	1.000	22.000	.001	.430	16.574	.973
Wilks' lambda	.570	16.574(b)	1.000	22.000	.001	.430	16.574	.973
Hotelling's trace	.753	16.574(b)	1.000	22.000	.001	.430	16.574	.973
Roy's largest root	.753	16.574(b)	1.000	22.000	.001	.430	16.574	.973

Each F tests the multivariate effect of DisplayType. These tests are based on the linearly independent pairwise comparisons among the estimated marginal means.

a Computed using alpha = .05

b Exact statistic

3. ManipulationType

Estimates

Measure: MEASURE_1

ManipulationType	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1	.723	.031	.658	.788
2	.554	.022	.509	.600
3	.863	.029	.803	.923
4	.906	.023	.859	.954
5	.942	.021	.899	.984
6	.967	.012	.943	.991

Pairwise Comparisons

Measure: MEASURE_1

(I) ManipulationType	(J) ManipulationType	Mean Difference (I-J)	Std. Error	Sig.(a)	95% Confidence Interval for Difference(a)	
					Lower Bound	Upper Bound
1	2	.168(*)	.033	.001	.060	.277
	3	-.140(*)	.039	.023	-.268	-.012
	4	-.183(*)	.038	.001	-.308	-.059
	5	-.219(*)	.029	.000	-.313	-.125
	6	-.244(*)	.030	.000	-.344	-.145
2	1	-.168(*)	.033	.001	-.277	-.060
	3	-.308(*)	.035	.000	-.424	-.193
	4	-.352(*)	.035	.000	-.466	-.238
	5	-.387(*)	.027	.000	-.477	-.298
	6	-.413(*)	.026	.000	-.497	-.328
3	1	.140(*)	.039	.023	.012	.268
	2	.308(*)	.035	.000	.193	.424
	4	-.043	.038	1.000	-.169	.082
	5	-.079	.029	.187	-.174	.017
	6	-.104(*)	.028	.017	-.196	-.012
4	1	.183(*)	.038	.001	.059	.308

	2	.352(*)	.035	.000	.238	.466
	3	.043	.038	1.000	-.082	.169
	5	-.035	.028	1.000	-.128	.057
	6	-.061	.024	.261	-.138	.017
5	1	.219(*)	.029	.000	.125	.313
	2	.387(*)	.027	.000	.298	.477
	3	.079	.029	.187	-.017	.174
	4	.035	.028	1.000	-.057	.128
	6	-.025	.017	1.000	-.081	.031
6	1	.244(*)	.030	.000	.145	.344
	2	.413(*)	.026	.000	.328	.497
	3	.104(*)	.028	.017	.012	.196
	4	.061	.024	.261	-.017	.138
	5	.025	.017	1.000	-.031	.081

Based on estimated marginal means

* The mean difference is significant at the .05 level.

a Adjustment for multiple comparisons: Bonferroni.

Multivariate Tests

	Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power(a)
Pillai's trace	.930	47.732(b)	5.000	18.000	.000	.930	238.660	1.000
Wilks' lambda	.070	47.732(b)	5.000	18.000	.000	.930	238.660	1.000
Hotelling's trace	13.259	47.732(b)	5.000	18.000	.000	.930	238.660	1.000
Roy's largest root	13.259	47.732(b)	5.000	18.000	.000	.930	238.660	1.000

Each F tests the multivariate effect of ManipulationType. These tests are based on the linearly independent pairwise comparisons among the estimated marginal means.

a Computed using alpha = .05
b Exact statistic

4. DisplayType * ManipulationType

Measure: MEASURE_1

DisplayType	ManipulationType	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
1	1	.902	.030	.839	.965
	2	.783	.033	.715	.850
	3	.798	.046	.704	.893
	4	.899	.039	.819	.979
	5	.985	.010	.964	1.006
	6	.993	.007	.977	1.008
2	1	.543	.054	.432	.655
	2	.326	.037	.250	.402
	3	.927	.042	.841	1.014
	4	.913	.038	.836	.991
	5	.898	.040	.815	.981
	6	.941	.020	.900	.983

Post Hoc Comparisons

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	SldHue	5526.4500	20	2242.06681	501.34138
	MLDHue	6339.5750	20	3781.38442	845.54326
Pair 2	SLDHoleAdd	1873.2500	20	655.10788	146.48657
	MLDHoleAdd	2275.3000	20	985.14292	220.28465
Pair 3	SLDHoleFill	1756.7750	20	469.30944	104.94078
	MLDHoleFill	2087.9250	20	1179.01086	263.63484
Pair 4	SLDGrow	6371.8250	20	2295.01515	513.18099
	MLDGrow	2838.3250	20	1642.58974	367.29423
Pair 5	SLDShrink	5673.3000	20	2419.16864	540.94255
	MLDShrink	1798.4500	20	562.64624	125.81152
Pair 6	SLDCorrectHUES	.9000	20	.14956	.03344
	MLDCorrectHUE	.6250	20	.15174	.03393
Pair 7	SXS_SHADE_NO RESP	.1250	20	.12825	.02868
	MLD_SHADE_NORESP	.0875	20	.14679	.03282
Pair 8	SLDCorrectHOLEADD	.9830	20	.05232	.01170
	MLDCorrectHOLEADD	.9410	20	.09835	.02199
Pair 9	SLDCorrectHOLEFILL	.9915	20	.03801	.00850
	MLDCorrectHOLEFILL	.9575	20	.07552	.01689
Pair 10	SLDCorrectSHADES	.7875	20	.16771	.03750
	MLDCorrectSHADE	.3375	20	.18629	.04166

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	SldHue & MLDHue	20	.462	.040
Pair 2	SLDHoleAdd & MLDHoleAdd	20	.050	.834
Pair 3	SLDHoleFill & MLDHoleFill	20	.402	.079
Pair 4	SLDGrow & MLDGrow	20	.039	.870
Pair 5	SLDShrink & MLDShrink	20	.136	.569
Pair 6	SLDCorrectHUES & MLDCorrectHUE	20	.145	.542
Pair 7	SXS_SHADE_NO RESP & MLD_SHADE_NORESP	20	-.262	.264
Pair 8	SLDCorrectHOLEADD & MLDCorrectHOLEADD	20	.090	.705
Pair 9	SLDCorrectHOLEFILL & MLDCorrectHOLEFILL	20	.397	.083
Pair 10	SLDCorrectSHADES & MLDCorrectSHADE	20	-.216	.361

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
					95% Confidence Interval of the Difference				
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper	t	df	Sig. (2-tailed)
Pair 1	SldHue - MLDHue	-813.12500	3388.89698	757.78040	-2399.17761	772.92761	-1.073	19	.297
Pair 2	SLDHoleAdd - MLDHoleAdd	-402.05000	1155.37715	258.35019	-942.78315	138.68315	-1.556	19	.136
Pair 3	SLDHoleFill - MLDHoleFill	-331.15000	1079.80470	241.45167	-836.51416	174.21416	-1.371	19	.186
Pair 4	SLDGrow - MLDGrow	3533.50000	2769.45178	619.26824	2237.35667	4829.64333	5.706	19	.000
Pair 5	SLDShrink - MLDShrink	3874.85000	2408.25133	538.50137	2747.75368	5001.94632	7.196	19	.000
Pair 6	SLDCorrectHUES - MLDCorrectHUE	.27500	.19702	.04405	.18279	.36721	6.242	19	.000
Pair 7	SXS_SHADE_NO RESP - MLD_SHADE_NORESP	.03750	.21877	.04892	-.06489	.13989	.767	19	.453
Pair 8	SLDCorrectHOLEADD - MLDCorrectHOLEADD	.04200	.10714	.02396	-.00815	.09215	1.753	19	.096
Pair 9	SLDCorrectHOLEFILL - MLDCorrectHOLEFILL	.03400	.06977	.01560	.00135	.06665	2.179	19	.042
Pair 10	SLDCorrectSHADES - MLDCorrectSHADE	.45000	.27625	.06177	.32071	.57929	7.285	19	.000

Appendix IX. Experiment 4: Repeated Measures ANOVA SPSS Tables

Experiment 4: Comparing Complex Stimuli

Response Times 2x3x4 Repeated Measures

Within-Subjects Factors

Measure:MEASURE_1

DisplayType	ManipulationType	StimuliType	Dependent Variable
1	1	1	SHADDRT
		2	CARADDRT
		3	GRPADDRT
		4	PHOADDRT
	2	1	SHDELRT
		2	CARDELRT
		3	GRPDELRT
		4	PHODELRT
	3	1	SHTRANSRT
		2	CARTRANSRT
		3	GRPTRANSRT
		4	PHOTRANSRT
2	1	1	SHADDRTMLD
		2	CARADDRTMLD
		3	GRPADDRTMLD
		4	PHOADDRTMLD
	2	1	SHDELRTMLD

	2	CARDELRTMLD
	3	GRPDELRTMLD
	4	PHODELRTMLD
3	1	SHTRANSRTMLD
	2	CARTRANSRTMLD
	3	GRPTRANSRTMLD
	4	PHOTRANSRTMLD

Descriptive Statistics

	Mean	Std. Deviation	N
SH-ADD-RT	5967.7941	2460.35441	17
CAR-ADD-RT	4571.0294	1843.17703	17
GRP-ADD-RT	3570.8529	908.18548	17
PHO-ADD-RT	7151.6471	3013.37996	17
SH-DEL-RT	4582.7059	1495.54940	17
CAR-DEL-RT	4201.0621	1459.51766	17
GRP-DEL-RT	3746.7353	1539.23876	17
PHO-DEL-RT	8024.9706	2502.43128	17
SH-TRANS-RT	7984.0000	3330.93733	17
CAR-TRANS-RT	8202.6176	3252.42338	17
GRP-TRANS-RT	4323.4118	1596.04457	17
PHO-TRANS-RT	8513.0000	2913.81055	17
SH-ADD-RT-MLD	4069.8529	1556.42146	17
CAR-ADD-RT-MLD	4124.3235	2177.78114	17

GRP-ADD-RT-MLD	4781.1176	1824.35483	17
PHO-ADD-RT-MLD	9406.4412	3804.42406	17
SH-DEL-RT-MLD	2847.6765	884.96788	17
CAR-DEL-RT-MLD	4624.4412	2849.28796	17
GRP-DEL-RT-MLD	4024.7941	1818.89049	17
PHO-DEL-RT-MLD	6066.5588	3522.90061	17
SH-TRANS-RT-MLD	2974.0588	1255.60145	17
CAR-TRANS-RT-MLD	4984.7647	3126.60736	17
GRP-TRANS-RT-MLD	3686.8529	1806.31083	17
PHO-TRANS-RT-MLD	6485.2647	2606.42359	17

Multivariate Tests[c]

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[b]
DisplayType	Pillai's Trace	.285	6.389[a]	1.000	16.000	.022	.285	6.389	.661
	Wilks' Lambda	.715	6.389[a]	1.000	16.000	.022	.285	6.389	.661
	Hotelling's Trace	.399	6.389[a]	1.000	16.000	.022	.285	6.389	.661
	Roy's Largest Root	.399	6.389[a]	1.000	16.000	.022	.285	6.389	.661
ManipulationType	Pillai's Trace	.513	7.906[a]	2.000	15.000	.005	.513	15.812	.904
	Wilks' Lambda	.487	7.906[a]	2.000	15.000	.005	.513	15.812	.904
	Hotelling's Trace	1.054	7.906[a]	2.000	15.000	.005	.513	15.812	.904
	Roy's Largest Root	1.054	7.906[a]	2.000	15.000	.005	.513	15.812	.904

StimuliType	Pillai's Trace	.901	42.302[a]	3.000	14.000	.000	.901	126.907	1.000
	Wilks' Lambda	.099	42.302[a]	3.000	14.000	.000	.901	126.907	1.000
	Hotelling's Trace	9.065	42.302[a]	3.000	14.000	.000	.901	126.907	1.000
	Roy's Largest Root	9.065	42.302[a]	3.000	14.000	.000	.901	126.907	1.000
DisplayType * ManipulationType	Pillai's Trace	.650	13.903[a]	2.000	15.000	.000	.650	27.805	.993
	Wilks' Lambda	.350	13.903[a]	2.000	15.000	.000	.650	27.805	.993
	Hotelling's Trace	1.854	13.903[a]	2.000	15.000	.000	.650	27.805	.993
	Roy's Largest Root	1.854	13.903[a]	2.000	15.000	.000	.650	27.805	.993
DisplayType * StimuliType	Pillai's Trace	.711	11.458[a]	3.000	14.000	.000	.711	34.374	.994
	Wilks' Lambda	.289	11.458[a]	3.000	14.000	.000	.711	34.374	.994
	Hotelling's Trace	2.455	11.458[a]	3.000	14.000	.000	.711	34.374	.994
	Roy's Largest Root	2.455	11.458[a]	3.000	14.000	.000	.711	34.374	.994
ManipulationType * StimuliType	Pillai's Trace	.800	7.339[a]	6.000	11.000	.002	.800	44.035	.981
	Wilks' Lambda	.200	7.339[a]	6.000	11.000	.002	.800	44.035	.981
	Hotelling's Trace	4.003	7.339[a]	6.000	11.000	.002	.800	44.035	.981
	Roy's Largest Root	4.003	7.339[a]	6.000	11.000	.002	.800	44.035	.981
DisplayType * ManipulationType * StimuliType	Pillai's Trace	.664	3.627[a]	6.000	11.000	.031	.664	21.763	.771
	Wilks' Lambda	.336	3.627[a]	6.000	11.000	.031	.664	21.763	.771
	Hotelling's Trace	1.978	3.627[a]	6.000	11.000	.031	.664	21.763	.771
	Roy's Largest Root	1.978	3.627[a]	6.000	11.000	.031	.664	21.763	.771

a. Exact statistic

b. Computed using alpha = .05

c. Design: Intercept

Within Subjects Design: DisplayType + ManipulationType + StimuliType + DisplayType * ManipulationType + DisplayType * StimuliType + ManipulationType * StimuliType + DisplayType * ManipulationType * StimuliType

Mauchly's Test of Sphericity[b]

Measure:MEASURE_1

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon[a]		
					Greenhouse-Geisser	Huynh-Feldt	Lower-bound
DisplayType	1.000	.000	0	.	1.000	1.000	1.000
ManipulationType	.890	1.746	2	.418	.901	1.000	.500
StimuliType	.756	4.114	5	.534	.864	1.000	.333
DisplayType * ManipulationType	.954	.708	2	.702	.956	1.000	.500
DisplayType * StimuliType	.812	3.062	5	.691	.902	1.000	.333
ManipulationType * StimuliType	.068	36.991	20	.013	.603	.801	.167
DisplayType * ManipulationType * StimuliType	.058	39.332	20	.007	.568	.741	.167

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

a. May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the Tests of Within-Subjects Effects table.

b. Design: Intercept

Within Subjects Design: DisplayType + ManipulationType + StimuliType + DisplayType * ManipulationType + DisplayType * StimuliType + ManipulationType * StimuliType + DisplayType * ManipulationType * StimuliType

Tests of Within-Subjects Effects

Measure:MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
DisplayType	Sphericity Assumed	1.154E8	1	1.154E8	6.389	.022	.285	6.389	.661
	Greenhouse-Geisser	1.154E8	1.000	1.154E8	6.389	.022	.285	6.389	.661
	Huynh-Feldt	1.154E8	1.000	1.154E8	6.389	.022	.285	6.389	.661
	Lower-bound	1.154E8	1.000	1.154E8	6.389	.022	.285	6.389	.661
Error(DisplayType)	Sphericity Assumed	2.890E8	16	1.806E7					
	Greenhouse-Geisser	2.890E8	16.000	1.806E7					
	Huynh-Feldt	2.890E8	16.000	1.806E7					
	Lower-bound	2.890E8	16.000	1.806E7					
ManipulationType	Sphericity Assumed	8.817E7	2	4.408E7	10.302	.000	.392	20.605	.979
	Greenhouse-Geisser	8.817E7	1.802	4.893E7	10.302	.001	.392	18.565	.969
	Huynh-Feldt	8.817E7	2.000	4.408E7	10.302	.000	.392	20.605	.979
	Lower-bound	8.817E7	1.000	8.817E7	10.302	.005	.392	10.302	.854
Error(ManipulationType)	Sphericity Assumed	1.369E8	32	4279071.921					
	Greenhouse-Geisser	1.369E8	28.833	4749135.105					
	Huynh-Feldt	1.369E8	32.000	4279071.921					
	Lower-bound	1.369E8	16.000	8558143.842					
StimuliType	Sphericity Assumed	7.434E8	3	2.478E8	49.556	.000	.756	148.669	1.000
	Greenhouse-Geisser	7.434E8	2.593	2.867E8	49.556	.000	.756	128.480	1.000
	Huynh-Feldt	7.434E8	3.000	2.478E8	49.556	.000	.756	148.669	1.000

	Lower-bound	7.434E8	1.000	7.434E8	49.556	.000	.756	49.556	1.000
Error(StimuliType)	Sphericity Assumed	2.400E8	48	5000307.941					
	Greenhouse-Geisser	2.400E8	41.482	5786067.984					
	Huynh-Feldt	2.400E8	48.000	5000307.941					
	Lower-bound	2.400E8	16.000	1.500E7					
DisplayType * ManipulationType	Sphericity Assumed	1.584E8	2	7.920E7	18.013	.000	.530	36.025	1.000
	Greenhouse-Geisser	1.584E8	1.912	8.285E7	18.013	.000	.530	34.438	1.000
	Huynh-Feldt	1.584E8	2.000	7.920E7	18.013	.000	.530	36.025	1.000
	Lower-bound	1.584E8	1.000	1.584E8	18.013	.001	.530	18.013	.978
Error(DisplayType*ManipulationType)	Sphericity Assumed	1.407E8	32	4396939.193					
	Greenhouse-Geisser	1.407E8	30.590	4599556.490					
	Huynh-Feldt	1.407E8	32.000	4396939.193					
	Lower-bound	1.407E8	16.000	8793878.386					
DisplayType * StimuliType	Sphericity Assumed	1.366E8	3	4.552E7	9.822	.000	.380	29.467	.996
	Greenhouse-Geisser	1.366E8	2.705	5.049E7	9.822	.000	.380	26.570	.993
	Huynh-Feldt	1.366E8	3.000	4.552E7	9.822	.000	.380	29.467	.996
	Lower-bound	1.366E8	1.000	1.366E8	9.822	.006	.380	9.822	.837
Error(DisplayType*StimuliType)	Sphericity Assumed	2.225E8	48	4634628.772					
	Greenhouse-Geisser	2.225E8	43.281	5139937.630					
	Huynh-Feldt	2.225E8	48.000	4634628.772					
	Lower-bound	2.225E8	16.000	1.390E7					
ManipulationType * StimuliType	Sphericity Assumed	1.078E8	6	1.796E7	4.355	.001	.214	26.131	.977
	Greenhouse-Geisser	1.078E8	3.620	2.978E7	4.355	.005	.214	15.765	.892
	Huynh-Feldt	1.078E8	4.809	2.241E7	4.355	.002	.214	20.944	.950
	Lower-bound	1.078E8	1.000	1.078E8	4.355	.053	.214	4.355	.501

Error(ManipulationType*StimuliType)	Sphericity Assumed	3.960E8	96	4124791.289					
	Greenhouse-Geisser	3.960E8	57.916	6837095.390					
	Huynh-Feldt	3.960E8	76.944	5146367.781					
	Lower-bound	3.960E8	16.000	2.475E7					
DisplayType * ManipulationType * StimuliType	Sphericity Assumed	7.774E7	6	1.296E7	3.498	.004	.179	20.985	.936
	Greenhouse-Geisser	7.774E7	3.410	2.280E7	3.498	.017	.179	11.927	.787
	Huynh-Feldt	7.774E7	4.446	1.749E7	3.498	.009	.179	15.550	.867
	Lower-bound	7.774E7	1.000	7.774E7	3.498	.080	.179	3.498	.420
Error(DisplayType*ManipulationType*StimuliType)	Sphericity Assumed	3.556E8	96	3704435.460					
	Greenhouse-Geisser	3.556E8	54.562	6517825.088					
	Huynh-Feldt	3.556E8	71.133	4999453.533					
	Lower-bound	3.556E8	16.000	2.223E7					

a. Computed using alpha = .05

Tests of Within-Subjects Contrasts

Measure:MEASURE_1

Source	DisplayType	ManipulationType	StimuliType	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
DisplayType	Linear			1.154E8	1	1.154E8	6.389	.022	.285	6.389	.661
Error(DisplayType)	Linear			2.890E8	16	1.806E7					
ManipulationType		Linear		1.310E7	1	1.310E7	3.240	.091	.168	3.240	.395
		Quadratic		7.507E7	1	7.507E7	16.624	.001	.510	16.624	.969
Error(ManipulationType)		Linear		6.468E7	16	4042221.995					
		Quadratic		7.225E7	16	4515921.847					
StimuliType			Linear	2.880E8	1	2.880E8	94.527	.000	.855	94.527	1.000
			Quadratic	2.620E8	1	2.620E8	36.657	.000	.696	36.657	1.000
			Cubic	1.934E8	1	1.934E8	40.228	.000	.715	40.228	1.000
Error(StimuliType)			Linear	4.875E7	16	3047109.652					
			Quadratic	1.144E8	16	7146981.825					
			Cubic	7.691E7	16	4806832.347					
DisplayType * ManipulationType	Linear	Linear		1.533E8	1	1.533E8	29.021	.000	.645	29.021	.999
		Quadratic		5081032.085	1	5081032.085	1.447	.246	.083	1.447	.205
Error(DisplayType*ManipulationType)	Linear	Linear		8.453E7	16	5283092.631					
		Quadratic		5.617E7	16	3510785.755					
DisplayType * StimuliType	Linear		Linear	8.732E7	1	8.732E7	19.706	.000	.552	19.706	.986
			Quadratic	4.516E7	1	4.516E7	7.800	.013	.328	7.800	.746
			Cubic	4081081.719	1	4081081.719	1.108	.308	.065	1.108	.168
Error(DisplayType*StimuliT	Linear		Linear	7.090E7	16	4431463.376					

ype)			Quadratic	9.264E7	16	5790232.472					
			Cubic	5.892E7	16	3682190.469					
ManipulationType * StimuliType		Linear	Linear	3.202E7	1	3.202E7	6.407	.022	.286	6.407	.662
			Quadratic	2.438E7	1	2.438E7	3.947	.064	.198	3.947	.463
			Cubic	3.071E7	1	3.071E7	10.797	.005	.403	10.797	.869
		Quadratic	Linear	9692220.862	1	9692220.862	3.834	.068	.193	3.834	.452
			Quadratic	7036509.423	1	7036509.423	1.321	.267	.076	1.321	.191
			Cubic	3958926.183	1	3958926.183	1.376	.258	.079	1.376	.197
Error(ManipulationType*StimuliType)		Linear	Linear	7.995E7	16	4997057.583					
			Quadratic	9.881E7	16	6175424.990					
			Cubic	4.551E7	16	2844106.344					
		Quadratic	Linear	4.045E7	16	2528280.382					
			Quadratic	8.524E7	16	5327499.107					
			Cubic	4.602E7	16	2876379.326					
DisplayType * ManipulationType * StimuliType	Linear	Linear	Linear	1422461.965	1	1422461.965	.308	.587	.019	.308	.082
			Quadratic	8191108.832	1	8191108.832	2.430	.139	.132	2.430	.311
			Cubic	3304628.353	1	3304628.353	.866	.366	.051	.866	.141
		Quadratic	Linear	5.269E7	1	5.269E7	13.560	.002	.459	13.560	.932
			Quadratic	9575884.275	1	9575884.275	4.838	.043	.232	4.838	.542
			Cubic	2554260.111	1	2554260.111	.561	.465	.034	.561	.109
Error(DisplayType*ManipulationType*StimuliType)	Linear	Linear	Linear	7.397E7	16	4622871.768					
			Quadratic	5.394E7	16	3371433.385					
			Cubic	6.105E7	16	3815667.132					
		Quadratic	Linear	6.217E7	16	3885794.075					
			Quadratic	3.167E7	16	1979472.909					
			Cubic	7.282E7	16	4551373.493					

a. Computed using alpha = .05

Tests of Between-Subjects Effects

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
Intercept	1.177E10	1	1.177E10	475.281	.000	.967	475.281	1.000
Error	3.963E8	16	2.477E7					

a. Computed using alpha = .05

Estimated Marginal Means

1. Grand Mean

Measure:MEASURE_1

Mean	Std. Error	95% Confidence Interval	
		Lower Bound	Upper Bound
5371.499	246.388	4849.179	5893.819

2. DisplayType

Estimates

Measure:MEASURE_1

DisplayType	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1	5903.319	288.394	5291.950	6514.687
2	4839.679	356.057	4084.872	5594.486

Pairwise Comparisons

Measure:MEASURE_1

(I) DisplayType	(J) DisplayType	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a]	
					Lower Bound	Upper Bound
1	2	1063.640*	420.793	.022	171.599	1955.681
2	1	-1063.640*	420.793	.022	-1955.681	-171.599

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

a. Adjustment for multiple comparisons: Bonferroni.

Multivariate Tests

	Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[b]
Pillai's trace	.285	6.389[a]	1.000	16.000	.022	.285	6.389	.661
Wilks' lambda	.715	6.389[a]	1.000	16.000	.022	.285	6.389	.661
Hotelling's trace	.399	6.389[a]	1.000	16.000	.022	.285	6.389	.661
Roy's largest root	.399	6.389[a]	1.000	16.000	.022	.285	6.389	.661

Each F tests the multivariate effect of DisplayType. These tests are based on the linearly independent pairwise comparisons among the estimated marginal means.

a. Exact statistic

b. Computed using alpha = .05

3. ManipulationType

Estimates

Measure:MEASURE_1

ManipulationType	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1	5455.382	286.208	4848.649	6062.116
2	4764.868	268.809	4195.019	5334.717
3	5894.246	301.457	5255.187	6533.306

Pairwise Comparisons

Measure:MEASURE_1

(I) ManipulationType	(J) ManipulationType	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a]	
					Lower Bound	Upper Bound
1	2	690.514*	216.473	.017	111.876	1269.152
	3	-438.864	243.812	.272	-1090.582	212.854
2	1	-690.514*	216.473	.017	-1269.152	-111.876
	3	-1129.378*	287.189	.004	-1897.045	-361.712
3	1	438.864	243.812	.272	-212.854	1090.582
	2	1129.378*	287.189	.004	361.712	1897.045

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

a. Adjustment for multiple comparisons: Bonferroni.

Multivariate Tests

	Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[b]
Pillai's trace	.513	7.906[a]	2.000	15.000	.005	.513	15.812	.904
Wilks' lambda	.487	7.906[a]	2.000	15.000	.005	.513	15.812	.904
Hotelling's trace	1.054	7.906[a]	2.000	15.000	.005	.513	15.812	.904
Roy's largest root	1.054	7.906[a]	2.000	15.000	.005	.513	15.812	.904

Each F tests the multivariate effect of ManipulationType. These tests are based on the linearly independent pairwise comparisons among the estimated marginal means.

a. Exact statistic

b. Computed using alpha = .05

4. StimuliType

Estimates

Measure:MEASURE_1

StimuliType	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1	4737.681	300.916	4099.769	5375.594
2	5118.040	342.818	4391.299	5844.781
3	4022.294	191.647	3616.022	4428.567
4	7607.980	380.912	6800.482	8415.479

Pairwise Comparisons

Measure:MEASURE_1

(I) StimuliType	(J) StimuliType	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a]	
					Lower Bound	Upper Bound
1	2	-380.358	361.838	1.000	-1468.888	708.171
	3	715.387	288.360	.148	-152.095	1582.870
	4	-2870.299*	274.472	.000	-3696.002	-2044.596
2	1	380.358	361.838	1.000	-708.171	1468.888
	3	1095.746*	280.471	.008	251.995	1939.496
	4	-2489.941*	327.797	.000	-3476.063	-1503.818
3	1	-715.387	288.360	.148	-1582.870	152.095
	2	-1095.746*	280.471	.008	-1939.496	-251.995
	4	-3585.686*	335.774	.000	-4595.805	-2575.567
4	1	2870.299*	274.472	.000	2044.596	3696.002
	2	2489.941*	327.797	.000	1503.818	3476.063
	3	3585.686*	335.774	.000	2575.567	4595.805

Based on estimated marginal means

a. Adjustment for multiple comparisons: Bonferroni.

*. The mean difference is significant at the .05 level.

Multivariate Tests

	Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[b]
Pillai's trace	.901	42.302[a]	3.000	14.000	.000	.901	126.907	1.000
Wilks' lambda	.099	42.302[a]	3.000	14.000	.000	.901	126.907	1.000
Hotelling's trace	9.065	42.302[a]	3.000	14.000	.000	.901	126.907	1.000
Roy's largest root	9.065	42.302[a]	3.000	14.000	.000	.901	126.907	1.000

Each F tests the multivariate effect of StimuliType. These tests are based on the linearly independent pairwise comparisons among the estimated marginal means.

a. Exact statistic

b. Computed using alpha = .05

5. DisplayType * ManipulationType

Measure:MEASURE_1

DisplayType	ManipulationType	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
1	1	5315.331	342.589	4589.075	6041.587
	2	5138.868	277.343	4550.927	5726.809
	3	7255.757	429.838	6344.541	8166.974
2	1	5595.434	406.129	4734.479	6456.389
	2	4390.868	408.027	3525.889	5255.846
	3	4532.735	417.787	3647.067	5418.404

6. DisplayType * StimuliType

Measure:MEASURE_1

DisplayType	StimuliType	Mean	Std. Error	95% Confidence Interval Lower Bound	Upper Bound
1	1	6178.167	500.499	5117.156	7239.177
	2	5658.236	347.407	4921.765	6394.707
	3	3880.333	231.966	3388.587	4372.080
	4	7896.539	496.502	6844.002	8949.076
2	1	3297.196	198.200	2877.031	3717.361
	2	4577.843	570.924	3367.538	5788.148
	3	4164.255	310.910	3505.154	4823.355
	4	7319.422	522.624	6211.509	8427.335

7. ManipulationType * StimuliType

Measure:MEASURE_1

ManipulationType	StimuliType	Mean	Std. Error	95% Confidence Interval Lower Bound	Upper Bound
1	1	5018.824	427.354	4112.874	5924.773
	2	4347.676	314.180	3681.644	5013.709
	3	4175.985	255.488	3634.375	4717.595
	4	8279.044	647.928	6905.498	9652.590
2	1	3715.191	210.625	3268.687	4161.696
	2	4412.752	372.017	3624.111	5201.393
	3	3885.765	260.081	3334.417	4437.112

	4	7045.765	541.406	5898.035	8193.494
3	1	5479.029	445.435	4534.749	6423.309
	2	6593.691	614.860	5290.246	7897.137
	3	4005.132	290.451	3389.404	4620.861
	4	7499.132	422.529	6603.412	8394.853

8. DisplayType * ManipulationType * StimuliType

Measure:MEASURE_1

DisplayType	ManipulationType	StimuliType	Mean	Std. Error	95% Confidence Interval	
					Lower Bound	Upper Bound
1	1	1	5967.794	596.724	4702.797	7232.792
		2	4571.029	447.036	3623.355	5518.704
		3	3570.853	220.267	3103.907	4037.799
		4	7151.647	730.852	5602.310	8700.984
	2	1	4582.706	362.724	3813.765	5351.646
		2	4201.062	353.985	3450.647	4951.477
		3	3746.735	373.320	2955.332	4538.139
		4	8024.971	606.929	6738.339	9311.602
	3	1	7984.000	807.871	6271.390	9696.610
		2	8202.618	788.829	6530.376	9874.859
		3	4323.412	387.098	3502.801	5144.022
		4	8513.000	706.703	7014.857	10011.143
2	1	1	4069.853	377.488	3269.615	4870.091

		2	4124.324	528.190	3004.612	5244.035
		3	4781.118	442.471	3843.121	5719.114
		4	9406.441	922.708	7450.387	11362.496
	2	1	2847.676	214.636	2392.668	3302.685
		2	4624.441	691.054	3159.472	6089.410
		3	4024.794	441.146	3089.607	4959.981
		4	6066.559	854.429	4255.250	7877.867
	3	1	2974.059	304.528	2328.488	3619.630
		2	4984.765	758.314	3377.212	6592.318
		3	3686.853	438.095	2758.134	4615.572
		4	6485.265	632.151	5145.165	7825.364

Accuracy 2*3*4 Repeated Measures

Within-Subjects Factors

Measure:MEASURE_1

DisplayType	ManipulationType	StimuliType	Dependent Variable
1	1	1	SHAddCorrect
		2	SLD_CartAddCorrect
		3	SLD_GrpAddCorrect
		4	SLD_PhotoAddCorrect
	2	1	SHDelCorrect
		2	SLD_CartDelCorrect
		3	SLD_GrpDelCorrect
		4	SLD_PhotodelCorrect
	3	1	SHTransCorrect
		2	SLD_CartTransCorrect
		3	SLD_GrpTransCorrect
		4	SLD_PhotoTransCorrect
2	1	1	MLD_SHAddCorrect
		2	MLD_CartAddCorrect
		3	MLD_GrpAddCorrect
		4	MLD_PhotoAddCorrect
	2	1	MLD_SHDelCorrect

	2	MLD_CartDelCorrect
	3	MLD_GrpDelCorrect
	4	MLD_PhotoDelCorrect
3	1	MLD_SHTransCorrect
	2	MLD_CartTransCorrect
	3	MLD_GrpTransCorrect
	4	MLD_PhotoTransCorrect

Descriptive Statistics

	Mean	Std. Deviation	N
SHAddCorrect	.8679	.19013	24
SLD_CartAddCorrect	.7917	.18158	24
SLD_GrpAddCorrect	.9167	.13077	24
SLD_PhotoAddCorrect	.7833	.17611	24
SHDelCorrect	.9025	.16979	24
SLD_CartDelCorrect	.6917	.42927	24
SLD_GrpDelCorrect	.8500	.18882	24
SLD_PhotodelCorrect	.8333	.18337	24
SHTransCorrect	.9579	.08895	24
SLD_CartTransCorrect	.4500	.18882	24
SLD_GrpTransCorrect	.9750	.08969	24
SLD_PhotoTransCorrect	.6583	.23204	24
MLD_SHAddCorrect	.8813	.15903	24

MLD_CartAddCorrect	.8333	.20990	24
MLD_GrpAddCorrect	.8167	.20359	24
MLD_PhotoAddCorrect	.4500	.19781	24
MLD_SHDelCorrect	.9504	.07893	24
MLD_CartDelCorrect	.7917	.21653	24
MLD_GrpDelCorrect	.9167	.13077	24
MLD_PhotoDelCorrect	.4583	.23204	24
MLD_SHTransCorrect	.9579	.08895	24
MLD_CartTransCorrect	.7667	.30455	24
MLD_GrpTransCorrect	.9750	.08969	24
MLD_PhotoTransCorrect	.6333	.26811	24

Multivariate Tests[c]

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[b]
DisplayType	Pillai's Trace	.048	1.162[a]	1.000	23.000	.292	.048	1.162	.178
	Wilks' Lambda	.952	1.162[a]	1.000	23.000	.292	.048	1.162	.178
	Hotelling's Trace	.051	1.162[a]	1.000	23.000	.292	.048	1.162	.178
	Roy's Largest Root	.051	1.162[a]	1.000	23.000	.292	.048	1.162	.178
ManipulationType	Pillai's Trace	.005	.051[a]	2.000	22.000	.951	.005	.101	.057
	Wilks' Lambda	.995	.051[a]	2.000	22.000	.951	.005	.101	.057
	Hotelling's Trace	.005	.051[a]	2.000	22.000	.951	.005	.101	.057
	Roy's Largest Root	.005	.051[a]	2.000	22.000	.951	.005	.101	.057

StimuliType	Pillai's Trace	.920	80.032[a]	3.000	21.000	.000	.920	240.097	1.000
	Wilks' Lambda	.080	80.032[a]	3.000	21.000	.000	.920	240.097	1.000
	Hotelling's Trace	11.433	80.032[a]	3.000	21.000	.000	.920	240.097	1.000
	Roy's Largest Root	11.433	80.032[a]	3.000	21.000	.000	.920	240.097	1.000
DisplayType * ManipulationType	Pillai's Trace	.347	5.846[a]	2.000	22.000	.009	.347	11.691	.822
	Wilks' Lambda	.653	5.846[a]	2.000	22.000	.009	.347	11.691	.822
	Hotelling's Trace	.531	5.846[a]	2.000	22.000	.009	.347	11.691	.822
	Roy's Largest Root	.531	5.846[a]	2.000	22.000	.009	.347	11.691	.822
DisplayType * StimuliType	Pillai's Trace	.683	15.068[a]	3.000	21.000	.000	.683	45.204	1.000
	Wilks' Lambda	.317	15.068[a]	3.000	21.000	.000	.683	45.204	1.000
	Hotelling's Trace	2.153	15.068[a]	3.000	21.000	.000	.683	45.204	1.000
	Roy's Largest Root	2.153	15.068[a]	3.000	21.000	.000	.683	45.204	1.000
ManipulationType * StimuliType	Pillai's Trace	.799	11.918[a]	6.000	18.000	.000	.799	71.507	1.000
	Wilks' Lambda	.201	11.918[a]	6.000	18.000	.000	.799	71.507	1.000
	Hotelling's Trace	3.973	11.918[a]	6.000	18.000	.000	.799	71.507	1.000
	Roy's Largest Root	3.973	11.918[a]	6.000	18.000	.000	.799	71.507	1.000
DisplayType * ManipulationType * StimuliType	Pillai's Trace	.525	3.311[a]	6.000	18.000	.022	.525	19.869	.820
	Wilks' Lambda	.475	3.311[a]	6.000	18.000	.022	.525	19.869	.820
	Hotelling's Trace	1.104	3.311[a]	6.000	18.000	.022	.525	19.869	.820
	Roy's Largest Root	1.104	3.311[a]	6.000	18.000	.022	.525	19.869	.820

a. Exact statistic

b. Computed using alpha = .05

c. Design: Intercept

Within Subjects Design: DisplayType + ManipulationType + StimuliType + DisplayType * ManipulationType + DisplayType * StimuliType + ManipulationType * StimuliType + DisplayType * ManipulationType * StimuliType

Mauchly's Test of Sphericity[b]

Measure:MEASURE_1

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon[a] Greenhouse-Geisser	Epsilon[a] Huynh-Feldt	Epsilon[a] Lower-bound
DisplayType	1.000	.000	0	.	1.000	1.000	1.000
ManipulationType	.994	.130	2	.937	.994	1.000	.500
StimuliType	.644	9.564	5	.089	.789	.885	.333
DisplayType * ManipulationType	.926	1.684	2	.431	.931	1.000	.500
DisplayType * StimuliType	.527	13.908	5	.016	.738	.819	.333
ManipulationType * StimuliType	.152	39.155	20	.007	.615	.747	.167
DisplayType * ManipulationType * StimuliType	.240	29.614	20	.079	.759	.969	.167

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

a. May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the Tests of Within-Subjects Effects table.

b. Design: Intercept

Within Subjects Design: DisplayType + ManipulationType + StimuliType + DisplayType * ManipulationType + DisplayType * StimuliType + ManipulationType * StimuliType + DisplayType * ManipulationType * StimuliType

Tests of Within-Subjects Effects

Measure:MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
DisplayType	Sphericity Assumed	.061	1	.061	1.162	.292	.048	1.162	.178
	Greenhouse-Geisser	.061	1.000	.061	1.162	.292	.048	1.162	.178
	Huynh-Feldt	.061	1.000	.061	1.162	.292	.048	1.162	.178
	Lower-bound	.061	1.000	.061	1.162	.292	.048	1.162	.178
Error(DisplayType)	Sphericity Assumed	1.208	23	.053					
	Greenhouse-Geisser	1.208	23.000	.053					
	Huynh-Feldt	1.208	23.000	.053					
	Lower-bound	1.208	23.000	.053					
ManipulationType	Sphericity Assumed	.004	2	.002	.049	.952	.002	.099	.057
	Greenhouse-Geisser	.004	1.988	.002	.049	.951	.002	.098	.057
	Huynh-Feldt	.004	2.000	.002	.049	.952	.002	.099	.057
	Lower-bound	.004	1.000	.004	.049	.826	.002	.049	.055
Error(ManipulationType)	Sphericity Assumed	2.062	46	.045					
	Greenhouse-Geisser	2.062	45.730	.045					
	Huynh-Feldt	2.062	46.000	.045					
	Lower-bound	2.062	23.000	.090					
StimuliType	Sphericity Assumed	8.514	3	2.838	64.857	.000	.738	194.572	1.000

	Greenhouse-Geisser	8.514	2.367	3.597	64.857	.000	.738	153.509	1.000
	Huynh-Feldt	8.514	2.656	3.205	64.857	.000	.738	172.272	1.000
	Lower-bound	8.514	1.000	8.514	64.857	.000	.738	64.857	1.000
Error(StimuliType)	Sphericity Assumed	3.019	69	.044					
	Greenhouse-Geisser	3.019	54.438	.055					
	Huynh-Feldt	3.019	61.092	.049					
	Lower-bound	3.019	23.000	.131					
DisplayType * ManipulationType	Sphericity Assumed	.701	2	.350	7.702	.001	.251	15.404	.935
	Greenhouse-Geisser	.701	1.863	.376	7.702	.002	.251	14.346	.922
	Huynh-Feldt	.701	2.000	.350	7.702	.001	.251	15.404	.935
	Lower-bound	.701	1.000	.701	7.702	.011	.251	7.702	.757
Error(DisplayType*ManipulationType)	Sphericity Assumed	2.093	46	.045					
	Greenhouse-Geisser	2.093	42.843	.049					
	Huynh-Feldt	2.093	46.000	.045					
	Lower-bound	2.093	23.000	.091					
DisplayType * StimuliType	Sphericity Assumed	2.950	3	.983	17.561	.000	.433	52.684	1.000
	Greenhouse-Geisser	2.950	2.213	1.333	17.561	.000	.433	38.855	1.000
	Huynh-Feldt	2.950	2.458	1.200	17.561	.000	.433	43.171	1.000
	Lower-bound	2.950	1.000	2.950	17.561	.000	.433	17.561	.980
Error(DisplayType*StimuliType)	Sphericity Assumed	3.863	69	.056					
	Greenhouse-Geisser	3.863	50.888	.076					
	Huynh-Feldt	3.863	56.540	.068					
	Lower-bound	3.863	23.000	.168					
ManipulationType * StimuliType	Sphericity Assumed	1.551	6	.259	9.359	.000	.289	56.154	1.000
	Greenhouse-Geisser	1.551	3.691	.420	9.359	.000	.289	34.548	.999
	Huynh-Feldt	1.551	4.485	.346	9.359	.000	.289	41.973	1.000
	Lower-bound	1.551	1.000	1.551	9.359	.006	.289	9.359	.834

Error(ManipulationType*StimuliType)	Sphericity Assumed	3.812	138	.028					
	Greenhouse-Geisser	3.812	84.903	.045					
	Huynh-Feldt	3.812	103.150	.037					
	Lower-bound	3.812	23.000	.166					
DisplayType * ManipulationType * StimuliType	Sphericity Assumed	.864	6	.144	5.802	.000	.201	34.813	.997
	Greenhouse-Geisser	.864	4.554	.190	5.802	.000	.201	26.423	.987
	Huynh-Feldt	.864	5.817	.149	5.802	.000	.201	33.750	.997
	Lower-bound	.864	1.000	.864	5.802	.024	.201	5.802	.636
Error(DisplayType*ManipulationType*StimuliType)	Sphericity Assumed	3.425	138	.025					
	Greenhouse-Geisser	3.425	104.741	.033					
	Huynh-Feldt	3.425	133.783	.026					
	Lower-bound	3.425	23.000	.149					

a. Computed using alpha = .05

Tests of Within-Subjects Contrasts

Measure:MEASURE_1

Source	DisplayType	ManipulationType	StimuliType	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
DisplayType	Linear			.061	1	.061	1.162	.292	.048	1.162	.178
Error(DisplayType)	Linear			1.208	23	.053					
ManipulationType		Linear		.002	1	.002	.036	.852	.002	.036	.054
		Quadratic		.003	1	.003	.064	.802	.003	.064	.057
Error(ManipulationType)		Linear		1.079	23	.047					
		Quadratic		.983	23	.043					
StimuliType			Linear	3.166	1	3.166	62.662	.000	.732	62.662	1.000
			Quadratic	.194	1	.194	5.510	.028	.193	5.510	.614
			Cubic	5.154	1	5.154	113.170	.000	.831	113.170	1.000
Error(StimuliType)			Linear	1.162	23	.051					

			Quadratic	.810	23	.035					
			Cubic	1.047	23	.046					
DisplayType * ManipulationType	Linear	Linear		.673	1	.673	12.118	.002	.345	12.118	.915
		Quadratic		.027	1	.027	.774	.388	.033	.774	.135
Error(DisplayType*ManipulationType)	Linear	Linear		1.278	23	.056					
		Quadratic		.815	23	.035					
DisplayType * StimuliType	Linear		Linear	1.654	1	1.654	44.961	.000	.662	44.961	1.000
			Quadratic	1.204	1	1.204	16.412	.000	.416	16.412	.972
			Cubic	.093	1	.093	1.601	.218	.065	1.601	.228
Error(DisplayType*StimuliType)	Linear		Linear	.846	23	.037					
			Quadratic	1.687	23	.073					
			Cubic	1.331	23	.058					
ManipulationType * StimuliType		Linear	Linear	.027	1	.027	1.187	.287	.049	1.187	.181
			Quadratic	.260	1	.260	18.483	.000	.446	18.483	.984
			Cubic	1.180	1	1.180	45.463	.000	.664	45.463	1.000
		Quadratic	Linear	.005	1	.005	.256	.618	.011	.256	.077
			Quadratic	.008	1	.008	.179	.677	.008	.179	.069
			Cubic	.071	1	.071	1.755	.198	.071	1.755	.246
Error(ManipulationType*StimuliType)		Linear	Linear	.523	23	.023					
			Quadratic	.324	23	.014					
			Cubic	.597	23	.026					
		Quadratic	Linear	.445	23	.019					
			Quadratic	.993	23	.043					
			Cubic	.930	23	.040					
DisplayType * ManipulationType * StimuliType	Linear	Linear	Linear	.187	1	.187	6.204	.020	.212	6.204	.665
			Quadratic	.010	1	.010	.429	.519	.018	.429	.096
			Cubic	.215	1	.215	8.399	.008	.267	8.399	.793

		Quadratic	Linear	.106	1	.106	7.326	.013	.242	7.326	.736
			Quadratic	.074	1	.074	4.075	.055	.151	4.075	.490
			Cubic	.272	1	.272	7.132	.014	.237	7.132	.725
Error(DisplayType*ManipulationType*StimuliType)	Linear	Linear	Linear	.694	23	.030					
			Quadratic	.514	23	.022					
			Cubic	.589	23	.026					
		Quadratic	Linear	.334	23	.015					
			Quadratic	.416	23	.018					
			Cubic	.877	23	.038					

a. Computed using alpha = .05

Tests of Between-Subjects Effects

Measure:MEASURE_1

Transformed Variable:Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
Intercept	365.176	1	365.176	3242.726	.000	.993	3242.726	1.000
Error	2.590	23	.113					

a. Computed using alpha = .05

Estimated Marginal Means

1. Grand Mean

Measure:MEASURE_1

Mean	Std. Error	95% Confidence Interval	
		Lower Bound	Upper Bound
.796	.014	.767	.825

2. DisplayType

Estimates

Measure:MEASURE_1

DisplayType	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1	.807	.017	.772	.841
2	.786	.017	.751	.821

Pairwise Comparisons

Measure:MEASURE_1

(I) DisplayType	(J) DisplayType	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a]	
					Lower Bound	Upper Bound
1	2	.021	.019	.292	-.019	.060
2	1	-.021	.019	.292	-.060	.019

Based on estimated marginal means

a. Adjustment for multiple comparisons: Bonferroni.

Multivariate Tests

	Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[b]
Pillai's trace	.048	1.162[a]	1.000	23.000	.292	.048	1.162	.178
Wilks' lambda	.952	1.162[a]	1.000	23.000	.292	.048	1.162	.178
Hotelling's trace	.051	1.162[a]	1.000	23.000	.292	.048	1.162	.178

Roy's largest root	.051	1.162[a]	1.000	23.000	.292	.048	1.162	.178

Each F tests the multivariate effect of DisplayType. These tests are based on the linearly independent pairwise comparisons among the estimated marginal means.

a. Exact statistic

b. Computed using alpha = .05

3. ManipulationType

Estimates

Measure:MEASURE_1

ManipulationType	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1	.793	.019	.753	.832
2	.799	.021	.756	.843
3	.797	.016	.765	.829

Pairwise Comparisons

Measure:MEASURE_1

(I) ManipulationType	(J) ManipulationType	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a]	
					Lower Bound	Upper Bound
1	2	-.007	.021	1.000	-.060	.047
	3	-.004	.022	1.000	-.061	.053
2	1	.007	.021	1.000	-.047	.060
	3	.003	.022	1.000	-.054	.059
3	1	.004	.022	1.000	-.053	.061

2	-.003	.022	1.000	-.059	.054

Based on estimated marginal means

a. Adjustment for multiple comparisons: Bonferroni.

Multivariate Tests

	Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[b]
Pillai's trace	.005	.051[a]	2.000	22.000	.951	.005	.101	.057
Wilks' lambda	.995	.051[a]	2.000	22.000	.951	.005	.101	.057
Hotelling's trace	.005	.051[a]	2.000	22.000	.951	.005	.101	.057
Roy's largest root	.005	.051[a]	2.000	22.000	.951	.005	.101	.057

Each F tests the multivariate effect of ManipulationType. These tests are based on the linearly independent pairwise comparisons among the estimated marginal means.

a. Exact statistic

b. Computed using alpha = .05

4. StimuliType

Estimates

Measure:MEASURE_1

StimuliType	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1	.920	.017	.884	.955
2	.721	.028	.663	.778
3	.908	.016	.876	.941

Estimates

Measure:MEASURE_1

StimuliType	Mean	Std. Error	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
1	.920	.017	.884	.955
2	.721	.028	.663	.778
3	.908	.016	.876	.941
4	.636	.020	.596	.677

Pairwise Comparisons

Measure:MEASURE_1

(I) StimuliType	(J) StimuliType	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a] Lower Bound	95% Confidence Interval for Difference[a] Upper Bound
1	2	.199*	.026	.000	.124	.273
	3	.011	.019	1.000	-.043	.065
	4	.284*	.025	.000	.212	.355
2	1	-.199*	.026	.000	-.273	-.124
	3	-.188*	.027	.000	-.265	-.110
	4	.085	.031	.068	-.004	.173
3	1	-.011	.019	1.000	-.065	.043
	2	.188*	.027	.000	.110	.265
	4	.272*	.019	.000	.218	.326

4	1	-.284*	.025	.000	-.355	-.212
	2	-.085	.031	.068	-.173	.004
	3	-.272*	.019	.000	-.326	-.218

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

a. Adjustment for multiple comparisons: Bonferroni.

Multivariate Tests

	Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[b]
Pillai's trace	.920	80.032[a]	3.000	21.000	.000	.920	240.097	1.000
Wilks' lambda	.080	80.032[a]	3.000	21.000	.000	.920	240.097	1.000
Hotelling's trace	11.433	80.032[a]	3.000	21.000	.000	.920	240.097	1.000
Roy's largest root	11.433	80.032[a]	3.000	21.000	.000	.920	240.097	1.000

Each F tests the multivariate effect of StimuliType. These tests are based on the linearly independent pairwise comparisons among the estimated marginal means.

a. Exact statistic

b. Computed using alpha = .05

5. DisplayType * ManipulationType

Measure:MEASURE_1

DisplayType	ManipulationType	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
1	1	.840	.022	.795	.885
	2	.819	.031	.754	.884
	3	.760	.018	.724	.797
2	1	.745	.025	.694	.797
	2	.779	.023	.732	.827

Pairwise Comparisons

Measure:MEASURE_1

(I) StimuliType	(J) StimuliType	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a] Lower Bound	Upper Bound
1	2	.199*	.026	.000	.124	.273
	3	.011	.019	1.000	-.043	.065
	4	.284*	.025	.000	.212	.355
2	1	-.199*	.026	.000	-.273	-.124
	3	-.188*	.027	.000	-.265	-.110
	4	.085	.031	.068	-.004	.173
3	1	-.011	.019	1.000	-.065	.043
	2	.188*	.027	.000	.110	.265
	4	.272*	.019	.000	.218	.326
4	1	-.284*	.025	.000	-.355	-.212
	2	-.085	.031	.068	-.173	.004
	3	-.272*	.019	.000	-.326	-.218

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

	3	.833	.026	.779	.888

6. DisplayType * StimuliType

Measure:MEASURE_1

DisplayType	StimuliType	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
1	1	.909	.024	.861	.958
	2	.644	.041	.559	.730
	3	.914	.021	.871	.957
	4	.758	.028	.700	.817
2	1	.930	.018	.894	.966
	2	.797	.038	.720	.875
	3	.903	.019	.864	.942
	4	.514	.030	.452	.576

7. ManipulationType * StimuliType

Measure:MEASURE_1

ManipulationType	StimuliType	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
1	1	.875	.025	.823	.926
	2	.813	.027	.756	.869
	3	.867	.029	.807	.926
	4	.617	.031	.552	.681
2	1	.926	.023	.879	.974

	2	.742	.050	.638	.846
	3	.883	.020	.843	.924
	4	.646	.032	.579	.713
3	1	.958	.018	.920	.995
	2	.608	.033	.539	.678
	3	.975	.018	.937	1.013
	4	.646	.031	.581	.711

8. DisplayType * ManipulationType * StimuliType

Measure:MEASURE_1

DisplayType	ManipulationType	StimuliType	Mean	Std. Error	95% Confidence Interval	
					Lower Bound	Upper Bound
1	1	1	.868	.039	.788	.948
		2	.792	.037	.715	.868
		3	.917	.027	.861	.972
		4	.783	.036	.709	.858
	2	1	.902	.035	.831	.974
		2	.692	.088	.510	.873
		3	.850	.039	.770	.930
		4	.833	.037	.756	.911
	3	1	.958	.018	.920	.995
		2	.450	.039	.370	.530

		3	.975	.018	.937	1.013
		4	.658	.047	.560	.756
2	1	1	.881	.032	.814	.948
		2	.833	.043	.745	.922
		3	.817	.042	.731	.903
		4	.450	.040	.366	.534
	2	1	.950	.016	.917	.984
		2	.792	.044	.700	.883
		3	.917	.027	.861	.972
		4	.458	.047	.360	.556
	3	1	.958	.018	.920	.995
		2	.767	.062	.638	.895
		3	.975	.018	.937	1.013
		4	.633	.055	.520	.747

Post Hoc Comparison Experiment 4

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	SLDTotal	4646.9792	24	742.12435	151.48550
	MLDTotal	3428.6042	24	1068.18941	218.04325
Pair 2	SLDTrans	6808.9792	24	2230.09421	455.21607
	MLDTrans	3598.4792	24	1321.69580	269.79003
Pair 3	SLDADD	5009.5833	24	1186.31632	242.15580
	MLDADD	4283.9792	24	1280.76661	261.43539
Pair 4	SLDDel	4605.9491	24	1011.46838	206.46512
	MLDDel	3723.4167	24	1310.49092	267.50284
Pair 5	MLDADD	4283.9792	24	1280.76661	261.43539
	MLDDel	3723.4167	24	1310.49092	267.50284
Pair 6	MLD PHOTOS RT	6694.8333	24	2381.17852	486.05603
	SXS PHOTOS RT	7268.3958	24	1862.72797	380.22775
Pair 7	MLD PHOTOS RT	6694.8333	24	2381.17852	486.05603
	SHAPES MLD RT TOTAL	2788.1458	24	992.33828	202.56020
Pair 8	SLDCorrect	.8262	24	.06385	.01303
	MLDCorrect	.7833	24	.08042	.01641
Pair 9	SLDCorrectTRANS	.6583	24	.10737	.02192
	MLDCorrectTRANS	.8183	24	.13470	.02750
Pair 10	SLDCorrectADD	.8392	24	.10794	.02203
	MLDCorrectADD	.7517	24	.11761	.02401
Pair 11	MLDCorrectADD	.7517	24	.11761	.02401
	MLDCorrectDEL	.7879	24	.10827	.02210
Pair 12	PhotoCorrectSLD	.7587	24	.13835	.02824

	PhotoCorrectMLD	.5146	24	.14575	.02975
Pair 13	PhotoCorrectMLD	.5146	24	.14575	.02975
	ShapeCorrectMLD	.8388	24	.09252	.01888

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	SLDTotal & MLDTotal	24	.178	.407
Pair 2	SLDTrans & MLDTrans	24	.021	.921
Pair 3	SLDADD & MLDADD	24	.086	.688
Pair 4	SLDDel & MLDDel	24	.312	.138
Pair 5	MLDADD & MLDDel	24	.558	.005
Pair 6	MLD PHOTOS RT & SXS PHOTOS RT	24	.083	.699
Pair 7	MLD PHOTOS RT & SHAPES MLD RT TOTAL	24	.480	.018
Pair 8	SLDCorrect & MLDCorrect	24	.271	.200
Pair 9	SLDCorrectTRANS & MLDCorrectTRANS	24	-.125	.560
Pair 10	SLDCorrectADD & MLDCorrectADD	24	.353	.091
Pair 11	MLDCorrectADD & MLDCorrectDEL	24	.432	.035
Pair 12	PhotoCorrectSLD & PhotoCorrectMLD	24	-.095	.660
Pair 13	PhotoCorrectMLD & ShapeCorrectMLD	24	.132	.538

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
					95% Confidence Interval of the Difference				
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper			
Pair 1	SLDTotal - MLDTotal	1218.37500	1187.58036	242.41383	716.90379	1719.84621	5.026	23	.000
Pair 2	SLDTrans - MLDTrans	3210.50000	2567.91718	524.17390	2126.16367	4294.83633	6.125	23	.000
Pair 3	SLDADD - MLDADD	725.60417	1668.87127	340.65692	20.90163	1430.30670	2.130	23	.044
Pair 4	SLDDel - MLDDel	882.53241	1383.60377	282.42694	298.28777	1466.77704	3.125	23	.005
Pair 5	MLDADD - MLDDel	560.56250	1217.87708	248.59812	46.29811	1074.82689	2.255	23	.034
Pair 6	MLD PHOTOS RT - SXS PHOTOS RT	-573.56250	2898.58697	591.67159	-1797.52843	650.40343	-.969	23	.342
Pair 7	MLD PHOTOS RT - SHAPES MLD RT TOTAL	3906.68750	2094.41163	427.51998	3022.29503	4791.07997	9.138	23	.000
Pair 8	SLDCorrect - MLDCorrect	.04292	.08809	.01798	.00572	.08012	2.387	23	.026
Pair 9	SLDCorrectTRANS - MLDCorrectTRANS	-.16000	.18247	.03725	-.23705	-.08295	-4.296	23	.000
Pair 10	SLDCorrectADD - MLDCorrectADD	.08750	.12858	.02625	.03321	.14179	3.334	23	.003
Pair 11	MLDCorrectADD - MLDCorrectDEL	-.03625	.12064	.02463	-.08719	.01469	-1.472	23	.155
Pair 12	PhotoCorrectSLD - PhotoCorrectMLD	.24417	.21024	.04291	.15539	.33294	5.690	23	.000
Pair 13	PhotoCorrectMLD - ShapeCorrectMLD	-.32417	.16197	.03306	-.39256	-.25577	-9.805	23	.000

Appendix X. Transcripts

TRANSCRIPT 1 – Island Gardens – Train Derailment – 04/04/2008

I: Tell me about the incident.

O: I received the action message was received at 6:15 am. It was an emergency call form the DLR manager and reported a railway accident. I deployed four units on a G2. The first unit deployed was a sergeant L22. The first update was about 06:40: L22 reported that there was "nothing suspicious, nothing terrorist related or criminal damage or vandalism. Somebody from the railway has left a piece of engineering equipment on the track. Traction current was off."

I: When you received the call, what were your specific goals?

O: Well, my goal was to send an officer to the scene to get a report. If there were passengers involved, there might be injured and wounded. It was necessary to get someone ASAP.

I: What features were you looking at when you received the call?

O: I was looking at my dispatch queue [the action queue] when I received the active message.

I: How did you know how many units to deploy?

O: An educated guess. It's a train derailment. L22 was a one-man, I knew more support will be needed, especially if there were passengers in the train. And there is the driver, someone had to talk/help him/her out…

I: What happened next?

O: The officer on scene reported that the driver had minor bruises, he was conscious and breathing and there was no bleeding. He said that there was nothing suspicious, but equipment has been left on the track. He requested more resources for crowd management.

I: What did you do next?

O: well, three more units were already on their way, but I sent 5 more officers and a duty sergeant. I also contacted RAIB because is a train related accident. Then another report came on the radio, there were 59 passengers; and the DLR manager and sergeant advised to shut the track down until midday. Top table told me to change the grade to G1. If officers go to a G1 (an immediate call), they go by with sirens but there are loads of things attached to that: possible low traffic injuries. So it's really important that we make sure the grade and the location are all right.

I: At any stage, were you uncertain about the *reliability or relevance of the information* that you used to formulate the decision? or about the *appropriateness* of your decision?

O: No, I was following procedures. When a call comes, one is never quite sure until one gets the first report.

I: What other courses of action were considered or were available to you? How was this option chosen or others rejected?

O: I was thinking about informing LAS but I decided to wait for the first report. And I was glad, there were no injured. LAS wasn't needed.

I: Are operators allowed to change the locations themselves?

O: Yes. But the supervisor might query why the location changed. There could be loads of reasons. If the original caller was heavily accented, unclear, and the operator genuinely believed that they heard correctly, that was the location, and someone makes another call to another operator, and confirms another location, then obviously, the location has to be changed. That could be just one reason.

I: You have only mentioned L22? Did the other units arrive?

O: [checking the incident log] This one didn't, but this one did, and…one unit was not logged as released. I think he never told us he arrived, and as a consequence he shows in the log he was never there. If the unit hasn't been released after 7 hours, he has never bothered to tell us. Or there might be a little story behind that. Or, the operator who received the call was trying to multitask, must have been busy, wrote it down somewhere, *All received thank you,* was in the middle of something, go ahead on another channel, and then forgot about it…

I: So what happened next?

O: I received another call reporting that traction current was remaining off at this time. I remembered that was a little bit worrying. Who turned it off? More important than that; who has responsibility for that? Whoever requested that to be turned off, is the only person that should be allowed to turn it back on. There are railway procedures and processes. So that was an interesting one. I checked for other updates in the action queue – I found out that the driver refused to blow the breathalyser. Pway confirmed that the equipment that caused the derailment was on permit [*Pway* stands for Permanent Way which a railway term for the track engineering crew]. The Pway track manager keeps track of large items like a generator. But the crew, they are obliged to walk the track before they released the track to the railway; from the start to the end of the engineering block. So the engineers take control on that line for a specified time say from 12 to 5 in the morning so there is no way a train would be able to get through. When the registry is handed it back to the railway, they HAVE to do a safety walk full length from the start to the end of the block. Since the equipment was still there, I called for heavy lifting gear to be sent to the location.

I: Were you at the time reminded of previous experiences in which a similar decision was made?

O: mmm, sort of. Last train derailment, we called SOCO (Scenes of Crime Officer) but SOCO don't go unless there is something to go to; there is no point, is it? They travel all over the south east of England. They wait for the first unit to arrive so the unit will confirm *Yes* we need SOCO or we don't.

I: So what happened next?

O: well, an update was received on the radio that three officers escorted the passengers to Deptford. No passengers were left on the train. CCTV was viewed and confirmed that equipment was left on the track the night before. Then we terminate our involvement and hand command over the railway people. And now that we stepped back, you see, now that is no longer ours, is now RAIB on the investigation because it is not criminal.

I: What do the "railway people" do?

O: The railway will do the PR stuff. In terms of service, there will be disruption to the line all day, and as a consequence that might end up in fried tempers - a little bit of pushing and shoving, and therefore, possible breaches of the law. The railway will put posters and press releases, local radio, TV news, depending on how serious it is. Say there was a murder, as a force, we would be closing that for 2 days, while we try to recover any evidence. In that case there will be a full blown, national press movement to ensure that everybody has been told. It doesn't matter if you send everyone in the country a personal letter, people will still turn up.

I: Why did you declare it critical if there were no injuries?

O: It is not a critical incident, trains derail. There is no passengers injured, there is no one left on the train. It is purely because of the effects on either the surrounding community or the political antenna. When we now declare something critical, we can change that any time, it means that everyone is made aware that there is something about this incident that we need to look closer, it could be anything from the reason why it happened, which is obviously very important; to the effect it had on the local community. That local community, for example, might be predominantly an ethnic minority group. In some cases, for example, if a black man has been pushed off the edge of the platform, or it was believed he was pushed off the edge, there could be a knock-on effect: are we going to have a little mini demonstration say at the front of the station on the police is racist all those kind of things, all those twitching antenna thing…

TRANSCRIPT 2

I: What features are you looking for when you receive a call?

O: The ladder tells you who is ringing and I double check the call ID sign.

I: Have there been any situations in which you missed some information?

O: As long as you are monitoring a radio call; the system records the conversation. So I don’t think I have missed anything, you can always rewind the conversation.

I: Is there anything that you will improve in the system to get more reliable information?

O: well yes, there is a need to improve the way to deploy resources and estimated time of arrival (eta). It becomes a problem when there are 10 units + running. When an officer calls, they have to confirm location, it will be great if automatically shows location in the log instead of us doing it manually. And a way to improve vehicles booking: if the police officer is not booked on a vehicle, the vehicle won’t be booked either.

I: Which is the mandated time for operators to deal with a call?

O: 90% of the call within 10 seconds, 90% of non-emergency calls within 30 seconds, 3 minutes to deploy units. In the report, the National calling handling standards will code the percentages in green if you have met the requirements, amber if the standards are intermediate or red if the standards haven’t been met.

I: What would jeopardize the call centre operations?

O. A communication failure - We need 10 minutes to change the system over

I: Do you use the map for locating the units? How do you know where are they?

O: The map is used for giving directions only, we can zoom in and have a very detailed view of streets and small roads; but not to locate units; the map gets updated every so often, but we don’t have time to sit there and watch it. It’s impractical. Units are supposed to update where they are all the time, so if they hit the emergency button, we will know.

I: Do you send units according to the grades?

O: We have to prioritize. Any incident graded as a disturbance comes as a *prompt*. And more than often we need to send officers. But if an ambulance is required, it comes up as *immediate*. But if someone faints you don’t need to send an officer. However, if urgent assistance is needed, you tell them, so you know everybody drops everything and do the urgent stuff.

You tell the top desk – we have 1 down, then you call an ambulance. If needed turn power off in the underground. And from there, others will take the necessary responsibilities.

I: Then, how do make sense of what’s happening? Did you use all the information available to you?

Is there any additional information that you might use to assist in the incident?

O: That’s why is important to have *Environmental Awareness* – when call comes in, you know by listening to the call taker that something is coming to you.

But sometimes, when you are really busy, these things are left out – just updating the duty status.

TRANSCRIPT 3

I: What do you look at as a call handler?

O: [She has other queues on the display so she displays only her AQ]. The call handler's action queue is the only queue that should be in. When a call is passed back to an operator, it dropped into that queue, and it would be red, and when is acknowledged it goes black. When they send it back, someone like me would pick it up, and when I take the action report, then I might need to send it back to the dispatchers; or it might not require any action and might need to go back into the AQ –

I: What features are you looking for when you take a call?

O: Anything that pops in there [pointing at the ICCS] is basically a job, and whoever is free picks it up. If it is urgent and it has come from the dispatchers [pointing at the action queue], for example someone is having a heart attack on Victoria, we'll shout *urgent request for ambulance in the AQ*. Not very technical but it does the trick. Of course, whoever is free, will hopefully look at that straight away, and do the necessary by contacting the ambulance service, tell them what is needed, getting the patient's number and then pops it back to the radio. One is flashing at the moment, is going red, because it has not been touched, it hadn't been picked up. If I was sitting in dispatch, anyone can pick up that call, whatever the incident: misper (missing person), travel card…

I: What is that camera for? [next to the call taker]

O. We can see who is coming in, that is the back entrance of the building; that is normally the side seen from upstairs, that one is the toilet at the back. But if anyone buzzes or rings up you can see who it is. Normally we cannot get CCTV, but that is because the terminal not working: … It is not our system; it is actually next door's system.

I: Who is next door?

Next door is the National Operation Rail. They are in charge of the whole underground, each individual line control like the district line control, the jubilee line control, each individual control room for the London Underground lines, and they got to see everything, they obviously see the whole lot.

I: If there is an incident in any of those lines, would you have access?

O. Yes, we have access to the CCTV. Have a look; there we did… a person under the train at Leicester square, like that… …

something like that [pointing at the screen] we can see the Fire brigade.

I: How does the system work?

O. Personal detained. … It is south bound in the Bakerloo line so we have a look on the other queues. the tablet points here, and then you have a look. Because it is not actually our system, it is next door's, it is used as a caveat. You get they have to come and ask for a personal detain we can have a look. So here, it comes back to the action queue [pointing to NSPIS]. So you can get an idea of what is going on. You will have a rough idea of what is going on

I: Then how do you make sense of what the situation is? [got interrupted by the chief inspector]

O: There are 7 major areas. You just get to know the areas, by training and practice.

TRANSCRIPT 4

I: What do you do as a call taker?

O. Basically you get a number of calls coming in with a telephone number, some come from the public. A car located, or they got pickpocket …or something like that. Not a lot of members of public know that if you get pickpocket in the underground, you should ring up BTP, they ring metropolitan Police…A lot of members of the public don't know that. But if the crime has happened on the underground or in the railway is BTP. So what happens, they call the Metropolitan Police, what they should do is log the incident and they pass it over on CAD. We check through, pick up the CAD, log in to NSPIS and it will go to the dispatch if it needs to.

I'll take calls mainly from out; a lot of them are just "Have you got telephone number for *these* please?" "the number of the nearest …", things like that. So we have got all that information, pretty much.

I: Are you supposed to be getting those phone calls?

O: They are not meant to come but, you know, all the numbers of police, fire brigade, ambulance, lost property… If you lose your bag in the underground that wasn't found so they think *lost property is my fault,* and they try to cancel a bank card, you are in the middle of London, you won't have your bank telephone number; we got all that as well, we have a list of all the telephone numbers here.

And then we'll get further calls from the train operators companies. Like southern rail or Clapham railway that their train has been … … Then it goes to dispatch, send officers there…

I: How do you determine the grade of the incident?

O: Ok, if it is a grade 1 that is when they go with the blue lights on. The computer automatically determines the grading which comes up here [depending on the channel it came through], but you can modify it.

So for example, we have been advised by railway staff there is person in the track between Clapham and … it has gone into grade 2. If they were babies, like toddlers, or a pregnant lady, if threats life or members of staff… then, say we have sent someone earlier, and a member of staff has been killed, we send grade 1

So for example, officer xxx just called me and said that she has just been assaulted, but the assailant is still there. So get it to the dispatcher, if the man is still there, so get officers get going, upgrade to grade 1. And once we got officers going, I'm still on the phone pulling up the description, because we have officers going now rather than pull a description and 5 minutes later get the officers going: So if it is a threat to life, we will change to grade 1. I forgot the actual terminology, I'm sure there is an official terminology for that.

I: What are the requirements in order to determine the grade of the call? How long does it take for units to arrive?

O. Grade 1 - Is immediate, usually injuries. Grade 2 – is a prompt. If is in London, it should be half an hour. Because the area we cover is so big, especially outside London, it could be an hour to an hour and a half. But, most of them will have station staff, but you should let them known. We also got routine which could be a call from the station which say they have found a bag here with large amount of cannabis, so "ok, you don't need to come and get it now, so I might come and pick it up later" so it will be a routine call. So you know most of the stations are aware that sometimes it takes time, especially out of London.

I: How do you find information in NSPIS?

O: We've got what we called *tagging list*. So, say you are sitting on radio, you go to your tea break and then you got something important or related incident, you come back here and link them. You can tag it all into one; it will sit in your "dispatch" tag queue, sit in the proper queue. Say I'm a radio operator, so I'll obviously be in the dispatch, I can't do a search for every incident because it has been quite busy at the moment, and I don't have a fantastic memory. So I have a look on the tags on the dispatcher queue just to be on the ground. Just to be tuned with dispatch. Then I can say "oh yeah, …" They [radio dispatchers] will create an incident and tag it. It is just easier looking things back (up). By looking things up, because everything goes back to the AQ [action queue]. If it is an ongoing incident, I created it, goes back to dispatch, once you have officers going, doesn't interfere, it is quite hard if it still sits in the available queue, than if you tagged it.

If you don't have the incident number, if it wasn't tagged, you won't be able to find it, you will have to search either for the call sign or for the actual location. It is quite easy to search but if you got loads of incidents queuing up … And once you finished, you are meant to clear it because it is done and dusted now.

When you have loads of things in the queue it [the tagging list] makes it quite easy. Because you can't see all that is in the action queue.

I: Can someone else go into your tag queue?

O: Yes, they can. Everything you'll do will show. If I tagged something, it will show. There's nothing you can do here that won't show. Some people tagged it to themselves, probably to follow up. I can untag anybody's. If I tag it, they untag it, it shows. This is used in the courts. All the information that goes on, everything can be printed off. Even if I changed male to female, it will show down here. You can't do anything without being logged, once it has been sent.

Once it has been closed, you can't do anything: It is locked. It is done and dusted. It is finished, is closed.

I: if you are making a decision on the steps to follow about a specific incident, what do you do – regarding the grading for instance?

O: Ok, if I have a call from the member of the public, who happened to be sitting in the carriage and someone is shouting and causing disturbance... He might be angry, he is shouting but he is not in any threat or attacking. So I'll send on a cop/call that would be Central London; if it is further up, I'll probably call the local police to help out instead.

If the member of the public rings back and says, she has gone mad, she is now biting my neck and kicking, and punched me in the knee; I'd upgrade that to G1. Log into NSPIS "kicking off, biting, make sure he is all right".

Or if you have got somebody holding from the edge of the bridge, grade 1 call. But if we have a missing person that maybe on that train then we don't send cops.

I: So the grading depends only on the type of incident?

O: yes and no. It sounds trivial, needless to say, but they [the officers] can see how the weather condition is like. If it is wet, if they are not able to drive, you can't do G1. They will get there when they get there, any particular time will be a record. If the incident is in London, it will probably be within 10 minutes,

If something is happening now, like a robbery, it goes as grade 1. For instance, this incident [pointing at the screen], I got the report from the security staff at the depot, the assailant was still there, so then I changed the grade to G1 because we want to catch him [the robber] red handed.

[Pointing at the screen, incident need to be acknowledge] This is not even our area; it is a job up north. The reason why it is coming through is just because it says network rail but is not one of ours, but the service is going on. There was a disturbance going on, this person has being abusive, but he is not being stupid. If he starts fighting, then the incident will get be upgraded.

So it depends, the systems grades the incidents automatically depending on the channel the call comes from, but then, one decides depending on the type of incident, and what the officers tell you. It is a little bit different.

I: In which situation would dispatch put something back to call takers? I don't understand why something in dispatch goes back to action queue…

O: Say they get called through, member of staff hit by a member of public. You hit a supervisor, say you didn't have money, and get a £20 fine. Police has gone, they arrive, they might look at the supervisor, he’s got a bleeding eye and a cut, and then they would call up and say could you call an ambulance please.

I: So they wouldn’t call an ambulance?

O. No, I would call an ambulance. What they would do, they’ll paste it on here, they will pass it back to the action queue, and they will probably shout across “LAS (London ambulance service) in the action queue” and then I’ll pick it up, have a look and I will ring the ambulance service. LAS need to know 4 main questions you always ask is: conscious, breathing, chest pain, still bleeding?

I: Once you created an incident, do you pass it to the dispatch queue? [1:30:38]

O: Only goes to the dispatch queue, if it requires police officers to attend. Sometimes it could be you are currently being assaulted in the underground, I will create an incident and transfer your call to our crime recording centre. The thief has just come out of your view, you got your wallet pick out of your pocket, you are not hurt, take a deep breath, I will pass you to our crime recording centre. I’m having a look, So for example, this gentleman has rang up, I have taken the details, I’ve taken his name, and telephone number, I take the name and telephone number first because if the line cut down I can’t called him back. I got his home address, the location. He rang me today but the actual attempt was discovered yesterday. So it is gone to the CRC queue, give them a crime record number and pass them to the CRC.

If the police are not required, you can pass the call to the front desk. A lot of people phone up for enquiries, or other police forces, but we don’t always create an incident. You might ring up and say that you need the phone for the American Embassy, so I won’t create an incident.

A lot of people call for information: we get a lot of admin, Oh I forgot the PCP for the underground station – they need to ring transport for London, What is the telephone number for the Polish Embassy? A lot of what we do, we shouldn’t be doing, they should be able to do it themselves.

If the police is required, pass to the dispatch radio, acknowledge incident, they will put the call out, and an officer will call them and say yes we can attend – eta 10 mins; or it will be half an hour, get the metropolitan police to help. So you have only 2 officers going they have 10 people, so we’ll pass over to CAD. The officers will on the radio give them updates.

Could require another action, depending on what the job is, so if you have something like a fatality, hundreds of people get called in. For example, the person under the train in Leicester Square, we have to let next door know. We called the fire brigade, ambulance service; we have to get our officers know, and you have the senior officers calling to find out what is going on, the met police coming so you got loads of information going backwards and forwards.

TRANSCRIPT 5 – Leicester Square – April 2008

I: I know you got the call from Leicester Square, could you please tell me what happened?

O: At 1600 the London Underground controller called about a person under a train in Leicester Square. Then hundreds of calls came from members of the public. I could see the active messages coming through.

I: When you received the call, what was the most important to accomplish at this point in the incident?

O: I graded as a G1. It was close to peak-time, central London. It was going to be crazy. The location is critical, and on Friday. So the first thing was to send help soon and inform the people next door: the Network Operation Rail. They are in charge of the whole underground.

I: What features were you looking for when you received the call? How did you know how many units to deploy?

O: Well, it is Leicester Square so that is always busy so we needed loads of officers and staff from the station for crowd control. A PCSO responded first, she was very close to the station, I send 5 PCSO and one officer. The officer reported back 8 minutes after the first call.

I: What did you do next?

O: I shouted across that we need LAS and the MET and the fire brigade. With all those people in Leicester… I also requested for the CCTV.

I: Did you check something else? What information did you use in making your decisions? What are the action lists that you use? How and where did you get this information, and from whom?

O: The first update came at 16:12 from the officer on scene. I always check the tagging list, constantly for updates. It's easier to find. The officer told us that there were about 3000 people in the station, but didn't know how many in the platform. An update in the tagging list informed us that there were 800 passengers in the train. Then I knew that we needed a lot of support, because we needed to talk with the people in the platform, see if it was a crime… So I sent 10 more officers and asked for the station staff to help with the evacuation of the station.

I: Was there any additional information that you might have used to assist in the formulation of your decision?

O: well it would have been great to know if it was a crime. But I advised RAIB (Rail Accident Investigation Branch) and CID (Crime Investigation Department).

TRANSCRIPT 6

CAD can cherry pick calls but no one else. No cherry pick is allowed.

10:28 Call transfer from MET

Check in NSPIS

10:30 "Can't find the job". Look for history. Check CAD

10:31 Start new log. Take notes on paper.

10:32 Finish call. Log incident in CAD

10:35 Transfer info to NSPIS

10:40 send it to front desk

In major priority emergencies, CAD operators shout to dispatchers the CAD log number so he/she can start action til CAD operator logs incident in NSPIS.

10:41 I'm stuck with Kent" Talk with the operator next to her.

10:41 Check CAD

10:42 "How do you pronounce the name"

Uche…

10:43 the operator called the person named Uche. Log in NSPIS. Where were you travelling? Type the address

10:44 forward to CRC (Crime Recording Center) by clicking radio system

10:45 the operators talks with CRC and put the person through – log action

10:59 CAD message arrive

11:05 Check Gazetteer

Assign unit

11:06 Call station

11:08 Dispatcher asked about incident 140

CAD operator shouts they have triangulated it

Dispatcher on phone explaining

CAD operators tells him to check by street "not too far"

11:09 type message to CAD MET

11:10 officer calls dispatcher

11:11 CAD operator types message in to CAD MET “No disturbance seen”

Closes log in NSPIS

I: What does the front desk do?

O: The front desk is an auditing quality control mechanism. It deals with incoming and outcoming faxes, copies of crime reports, and paper records but those are kept for 31 days then shredded. It is the POC (point of contact) for PNC (Police National Call) broadcast.

The front desk is also in charge of the station check system – it comes once an hour, the system automatically checks and prints a report of all the stations status. Only mainland rail have to call to update the status. It is mainly for terrorism check

CAT1 – GENUINE THREAT / CAT2 – MAY NOT BE GENUINE.

TRANSCRIPT 7

12:17 types call id BQ##. Finds the log and the incident number

12:18 “all received”. Add log

12:19 assigns units in the system

12:20 call B49

12:21 Releases units. Write down id’s in paper

12:23 BQXX calls informing where the prisoner will be taken

Logs info.

12:24 responds call BJ606 asks for incident umber

12:32 TD164 trainee (operator takes note of call sign)

12:32 TD160 is at Paddington (operator takes note of call sign) “Yes yes so received”

12:33 update call sign in NSPIS

12:41 Operator asks CAD to double check spelling of the name [shouting across the room]

12:42 assign a unit to the event

12:52 stands up to update the whiteboard (indicates who the duty officer is in each area)

12:52 comes back acknowledge message in NSPIS. Contact the officer.

I: Do you use the map?

O: The mapping system is not so good. For instance it doesn’t differentiate the date. It is the national mapping system but it is not so user-friendly. We use the map for example to find addresses, but you can do that through Gazetteer which has live access to the station rail system.

I: What do you look at when you are working with the NSPIS? How do you know that something is important?

O: It’s all colour coded. If the message is red, it means that it hasn’t been acknowledge so anyone who is free has to acknowledge it. If it is green, it means that it hasn’t been resourced yet. If it turns blue, it means that someone has updated the incident log. Finally all these black messages, they have been acknowledged and resources have been deployed.

I: and the ICCS?

O: The ladder gives you the call id. When the call comes up, here [pointing the channel] it gives you the radio sign too. The incoming calls on this side [pointing the right-hand side of the screen] tells you who is calling, say LU for London Underground or B56, so I know is the duty sergeant. If an emergency comes up, everything flashes, the call sign should come up in the system, if not; the operator has to challenge the officer to identify him or herself for alias.

I: What are the incidents' grades?

The grades of the incidents are:

- Immediate
- Prompt
- Routine
- Police generated
- Deferred
- Non-attendance
- Crime recording

I: What does the PNC checks?

- Persons – if they have any criminal record
- Gun licenses
- Vehicles – if they have been declared lost or stolen
- Driver licence – validity
- Property index
- Valuable animals

TRANSCRIPT 8 – Liverpool Street Station 7/7/2005 (7.5 years of experience)

I: During your time here in the control room, there must have been some incidents that you cannot forget. I would like you to tell me about a critical incident where you were interrupted, or one that escalated after a break.

O: Oh there is one. The day of the London Bombings. The first call I received was a power outage at Liverpool Station at 8:51. There first update was a PCSO at Liverpool station stating there was dust or smoke. I told him to take a look if he can. So he came back with an update: a member of the staff said they had a power failure.

I: Then, what did you do?

O: First, I informed the duty officer about the power failure. He was already aware. Then, I was trying to establish what happened: At Liverpool Street Station you have many lines, was it one line either directions, or all lines affected?

I: Did you check any of the action queues for more information?

O: I didn't have to, seconds after, a second call came in: person under the train Edgware Road Station. I dispatched 2 or 3 units. I can't remember. Then, I got the first report from the unit on scene. The officer was downstairs – "It could be something else dust and smoke coming out".

I went to talk with the supervisor and overheard Bomb damage.

I ran back to the radio – called the unit in place and asked him to confirm bomb damage. He confirmed it was suspicious.

I: So what happened?

O: I sent more units to both Liverpool Street Station and Edgware road. I needed to check if it was linked. Then the senior officers from upstairs (LU) came down. Everybody was talking to me at the same time. They told me to change channels. I became AZ, the link between Tavistock FCRL and the Glasgow respond unit. They told me there was a swap NO G8, we take incidents in London. [We were monitoring the G8 Summit]. It got very confusing; everybody was talking at the same time.

I: as an AZ, what else did you have to do?

O: I was supposed to monitor the hand over of the G8, but everything was crazy. I then was asked to tell all bronze officers that a telephone conference was on place. I put the call out to all bronze officers. A senior officer told me then that they might not know who the bronze officers are. But with so much radio traffic, it wasn't my concern. I can't help if they don't know.

I: Were you keeping track of what was happening? What are the action lists were you using?

O: While I was in the radio with Glasgow, I was still log in the action queue monitoring the messages. I saw a call came in from Russell Square, this was different. The first report said “multiple injured, limbs missing, people covered in dust”; second report “other emergency services were in route”. I overhead the top table was getting information from the NOC (Network Operation Centre).

Then, AZ [alfa zulu] was ready for hand over. I asked them; they said yes, I came back to my channel in London. I checked the tagging list straight away, but I knew what came from Russell Square.

I: Where there any changes that you could have missed while you were an alfa zulu?

O: No, As I said, when I was assigned to be an AZ, I was logged in both queues. Just click this button. I couldn’t miss a change. I knew what came from Russell Square…

I: Then, what happened?

O: A radio call from an officer came in, a bus exploded at Tavistock Square. I put a call out to all units. We didn’t have more units. BTP vehicle was on scene. Second update “Doctors from the British Medical association were providing immediate help”. I can’t remember anything else. I got back home at 20:00.

I: At any stage, were you uncertain about the reliability or relevance of the information that you had?

Or were you uncertain about the appropriateness of your decision?

O: At the beginning it was chaos. First, Liverpool Station, I didn’t know how many lines were affected. Then, Edgware Road, and then straight after Russell Square. And the people from upstairs. It took more than 20 minutes to realize what was happening. Then we didn’t have any more units left. We were just logging the updates that came from the radio.

TRANSCRIPT 9 - Gainsborough Road, E11 - 08/03/2009 12:10 (5 years of experience)

I: Could you please think of an incident that has been extremely difficult to forget. Could you think back to one of those incidents that involved a significant amount of resource allocation?

O: There was a POLAC. This was a police accident. I: Tell me what happened?

O: I got a call on the radio and said *"POLAC"* and that's all they said.

I: So how did you know who was calling you? Or what happened?

O: Oh here, in the ICCS, it gives you the call sign. So I knew it was one of our underground vehicles going to Canary Wharf. The officer was assigned to a G1. So I knew where he was and where he was going to, but not the exact location.

I: What was the most important to accomplish at this point in the incident?

O: well, we knew he had been involved in an accident, I put a call out to all units "urgent assistance" which will get numerous units attending. We had 4 units attending. The first units that replied were deployed; that was 3 minutes after the call. One was a special response unit. I shouted over the room LAS required Gainsborough Road, POLAC.

I: Did you have any other additional information?

O: Not at this stage, I called the officer back and I managed to get the location of where he was from the officer that was driving the vehicle. And then his transmission stopped. So, we knew where he was.

I: Did you have to inform the supervisor? If yes, when did you do it?

O: As soon as I heard POLAC. I shouted over to the duty sergeant 159 POLAC and he was monitoring it, plugged into the radio and listened to the transmission. Then I shouted across LAS 159 POLAC to the call takers.

Interrupted…

I: So you got a call over the radio saying POLAC. You assigned 4 units, got an ambulance and informed the duty officer. Then what happened?

O: I still tried to call up the vehicle that was involved. I managed to get a response out of him asking for ambulance and the Metropolitan Police Traffic Unit. I passed it over to the call handler; I put it in the AQ. The request for ambulance was at

12:15. LAS had already a call from a member of the public. They were on their way. We informed the MET Traffic Unit (In the case of police road accident they will definitely have to go because the accident happened in their area, and they would photograph the scene, measure the distance; although it was involving one of ours, we will hand over the primacy of the investigation and they would be involved with the vehicle examination).

I: Did you get a report?

O: Six minutes from the call coming in the first unit arrived. Then we could get an update from an officer that was not involved in the accident; the report was "Multiple casualties". The officer in scene asked for Helicopter Ambulance due to critical injuries. "Officer Unconscious trapped in the car" he said. I put in the AQ to update the ambulance crew so it will save a few seconds on arrival multiple casualties and we also call the Fire Brigade.

I then cancelled the special unit but they were almost there. I redeployed them.

I: Did you use some other queue to get more information from, or just the officer on scene? How and where did you get information, and from whom? What did you do with the information?

O: well, I was checking the tagging list constantly. So I remembered that 11 minutes after the call came in, the other emergency services arrived to the scene so the MET, LAS and the Fire Brigade.

I asked for the Fire Brigade just as a precaution, because the officer was unconscious still in the vehicle. I didn't know how damaged the vehicle was, that was my primary concern at the time. Knowing how traffic accidents and ambulances, getting people out of cars, they would have needed the fire brigade anyway. They were already aware.

The officer that was assisting asked for the area duty officer, but we already asked him to go. Because of the seriousness, every police accident has to have a supervisor so a sergeant or above. The first officer that arrived he is a sergeant already. We also called the area duty officer to be pre-empted thinking that they may be needed in scene. We called the duty officer 15 seconds after LE5 was deployed.

We had another unit that arrived as well, 10 minutes after the initial call. They called up and offered us, they heard the developments over the radio, and they offered us instead of us calling and deploying them. They arrived.

I remembered a message from the CAD indicating that they were contacting the helicopter, and they could give us a landing point.

I went back to the CAD it was 12:23 when the message about requiring a helicopter –immediately- came in.

I: What happened next?

O: Then it was important to get an update from the officer on scene from the ambulance service. I called the first unit LE5, for an update. He supplied us with an update; he told me that the car has to be cut off. Possible neck and back injuries.

While this is still going on we had updates from the CAD system. Because the ambulance service will talk to the CAD system and passes it back to us. They have the same system. Ambulance said that the helicopter was not required at that time.

The special response unit A80 arrived. They have a camera on board. They were taking photographs of the cars, the final position of the cars, and the direction from where the vehicles came from, damages of the vehicles.

I: Did they come with an update?

O: No, they would have been dealing with the scene itself. At the same time, 5 LAS arrived. LE5 was dealing with the ambulances. He was our eyes and ears.

I: Did LE5 tell you about the ambulances?

O: Yes, but the message came up also in CAD. I was checking both.

I: Do you think you missed any information that could become relevant later?

O: No, I don’t think so. But I prompted the officer to provide with the pieces that were missing. LE5 came back, helicopter was not required. Ambulance Service had sufficient people there to deal with the ongoing problem, the injuries were not serious enough to warrant a helicopter.

I called LE5 again for vehicles registrations: 3 vehicles involved. Put in AQ for a PNC. Only here we knew that there were 3 vehicles involved including the police vehicle. Just in case, we have the details of the registry owners in case we needed it.

I called LE5 again for update on injured of the other cars. I needed to know where the officer was going to be taken to, and the other injured. The driver of one of the vehicles was treated of shock.

I: Did you get any additional information from any other source?

O: Yes, the superintendent reported that he was waiting for LA95 to arrive (sergeant), to attend and assess a full update. This was escalating higher up quite rapidly. LA95 was going with the injured officer. He would have been given a brief and then he rang the duty office with a brief incident summary of what was happening, their action plan, and any persons that they need to call. There was a member of the public in the pavement.

Then, I received a call from CID, they have been informed by I don’t know who and were going to the hospital. The CID will go to hospital to get any report of possible criminal negligence.

I also checked the CAD, and 35 minutes after the call came in, the Met Traffic Unit arrived. While all this is going on, the Met Police has to be looking at road closures, the buses will have to be told.

Now everything has calmed down, we have officers on scene dealing with the incident, officers dealing with road closures, duty sergeant dealing with next of kin, so we got very quiet here. We are just waiting for more updates. So that is the end of our involvement in the radio.

Was that the end of the incident?

O: We were waiting for updates. That is the end for the radio, but there are loads to be dealt with. There will be press enquiries that will be dealt by the call handlers. The duty officer will update a very brief press release so the call handlers can give that to the media people. The other impact we have to think about is the next of kin. I got a call from the sergeant LA95, he wanted to know the details about the next of kin. I got the details: home address and phone number, passed it to the duty officer to pass it on to the sergeant.

I: Do you think that you got irrelevant information?

O: Not so much irrelevant data, but stuff can get duplicated. You might be told something by the duty officer and get another called by the officer saying the same thing, or several forces informing us about the same incident.

You do get that in all incidents, there is a certain amount of duplication. It doesn’t matter if they have said it 3, 4 or 5 times. It is better to have it 5 times than not having.

I: What information did you have available after the called got interrupted?

O: Just the *Active Message - “POLAC”* , the call ID and that’s all.

I: What features were you looking for when the transmission stopped? How did you know that you needed that information to resume your task?

O: The location of the accident so we can send assistance. I managed to get the location of where it was from the officer that was driving the vehicle but then his transmission stopped.

I: What was the most important piece of information that you used to resume your task?

O: Location and the call ID.

I: What are the action lists that you use? How and where did you get this information, and from whom? What did you do with the information?

O: The officer who was driving gave me the location of the incident. The ID comes up when he called, so that's ICCS.

I: Is there any "rule of thumb" that you have or recommend?

O: There are procedures dealing with any incidents regardless of the nature. You get through the same motion, if you get an urgent assistance, or a routine, can you create an incident?

You usually know by the officer voice how urgent it is.

The only way that the job could be done better is if we could all type properly. That will obviously increase our response by getting the information into the log. There are very few people that can touch-type.

I: How do you create awareness of the situation?

O: It is just something that develops. I keep listening to both the radio and the environment. When you are first on the radio, you hear something on the radio but you won't hear anything else in the room, not even the person sitting next to you training you because you are concentrating so hard. But after some time, you can listen to the radio, type what they are saying, and listen to the call handler talking about the incident that you are dealing with, and look up. But you can have a conversation and hear the radio and hear to the other person.

I: When you receive a call, what do you imagine?

O: Normally I imagined what is happening there, then, I make a picture in my head. In this case, when I heard POLAC, I thought he hit someone on the road. When I lost transmission, I knew we needed more ambulances. Then that picture changed. I only knew how many vehicles after the 4th report - 3 vehicles

I: Does this first picture affect your actions?

O: Not sure, I don't think it wouldn't affect the actions; it would affect the urgency of my actions

I: Would you have made something differently if you would have had more information?

O: What could I have done differently, I always try to be self-critical. But in this case, there was nothing that could have been done differently apart from not happening. Because it was me that called him up to assist the Metropolitan Police. The call was to Canary Wharf. That is where he was asked to go to. But all ended well, he was fine when he got to hospital.

TRANSCRIPT 10 – London Bombings 7/7/2005 (8 years of experience)

I: During your time here in the control room, there must have been some incidents that you cannot forget. I would like you to tell me about a critical incident where you were interrupted, or one that escalated after a break.

O: the 7/7. It was early morning. I was assigned to “Radio relief” that day. Channel 2 [the Underground channel] went down, it ceased to exist, but it always goes down. So I went for a cup of tea and decided to sort that out later. We only had four radio operators at the time. We had two dispatchers in the main channel, and the radio relief, and somebody covering channel 2 which was the Underground channel.

Then, someone shouted “Smoke report at Liverpool Street Station”. So I thought, ok, we will get someone there to investigate.

I: What features were you looking for when you resume your task? How did you know that you needed that information to resume your task?

O: I came back and logged into the main channel. I was dealing with a ticket incident and then I had to prioritize I had loads of active messages. Active message active message active message - You want to keep the airwaves clear.

Active Message Edgware road “Numerous walking wounded, lost of lives and limbs”. But at the same time at Russell square “Walking wounded”.

I: Could you prioritise which piece of information is most important when resuming the task?

O: Because they were all the same, what I did, the first unit with the active message go ahead with your message, then second unit. They were asking for ambulances, for fire brigade.

Then I checked the tagging list and back to my dispatch queue: another active message “Walking wounded”. I have all these active messages. They are underground; their active message is active to them. Everybody else was doing the same thing, for them they were active messages, they were urgent; they were trying to get it across.

I wanted to prioritize the most important, so I said first unit stand by, second unit stand by, but they were all the same. So I did it in sequential order.

I: How did you make sense of what was going on?

O: You could get geared up for a critical incident happened at one location, but for several critical incidents happening all at the same time, it’s just unheard of. It’s confusing. It blows your mind. What’s going on? I thought.

I passed it to the top table. We have these, so we need units here, and here and here. So which units can I send?

I: While you were talking to your supervisor, did you miss any changes about the incident?

O: Well…No, I don’t think I missed anything. As soon as I came back I went to the tagging list. At the same time, I was on the radio sending out the units. I knew what was happening, at least from Liverpool Station…

I: Did you have a picture of what was going on?

O: No, I couldn’t, I didn’t. When I came back, someone said there’s been a one under at Edgware Road. So that’s all I have to go and looked it up. First it came as a person under a train. So I checked the one under at Edgware Road. But I have all these active messages about walking wounded, numerous multiple casualties, there are limbs missing. But so how can it be, did the train came out of the track, was it a train wreck? So I couldn’t understand what was going on. And then King’s Cross, Leicester Square, Liverpool Street. The initial point was total confusion.

I: What was the most important thing to accomplish at this point in the incident?

O: We were getting people to the locations but we didn’t know what was going on. Then we had problems with the radio because everyone wanted to talk at the same time. It took us 20 minutes. It seemed like forever. It took us a long time.

I: Do you think you could develop a rule, based on your experience, to help another person to make a successful decision? A rule of thumb, a tip?

O: Integrate. You have to keep a running log, because they are all talking to you. Take a paper, draw a grid, and go in sequence, if you can’t type, write everything down: call signs, descriptions. And then you type it in. Take charge of the situation, don’t be overruled.

I: What are the action lists that you use when resuming your task?

O: As I said, went back to the main channel. Always look the tagging list, G1 and G2. But in this case, they were all G1.

TRANSCRIPT 11 – Three Locks Bridge 2009

I: During your time here in the control room, there must have been some incidents that you cannot forget. Could you think back to one of those incidents that involved a significant amount of resource allocation?

O: There was this one incident close to Leighton Buzzard. A female officer at Euston called upset. I've heard about a person under a train at Leighton Buzzard. "Are you aware of anything?" "No we haven't been advised of anything! What's your information?"

She said that a train has been cancelled. Passengers were asking about the train that has been cancelled, if partly it's due with the person under the train at Leighton Buzzard. And that's all the details we had. All the details she had.

I: Did you do anything with that information?

O: Well, in the meantime, there were 3 units that called at the same time to be assigned. I'm on route one zero, I'm on route eta two zero … I said "hang on; we don't know if anything has happened". So I picked one unit, the BG [bravo golf] unit, but only on a G2. Because at the moment it is not a G1 on the information we got, because it's unconfirmed. So I am not going to send units in their blues and twos for nothing.

I: What was the most important thing to accomplish at this point in the incident?

O: Well, it was to enquire, to confirm the incident. Because my thinking was why would passengers at Euston Station know of a person under a train at Leighton Buzzard 40 miles away before we did? And that was my reasoning. So I cancelled the other 2 units. I send only one unit to go to Leighton just to make enquiries.

I: What happened next?

O: I heard Bryan X speaking "Does anybody know where such and such bridge is?" Because I knew, I am from up there. My ears picked up at the sound of this bridge. So what have you got? "I don't know, I just had Network Rail Birmingham on telling me about an incident a person under a train." Ok, thank you Bryan, ok.

So I went back to the radio. "Units now, now you can go" I get the units running again. Now it's confirmed. So I said "Three Locks Bridge", not Leighton Buzzard. It was between Leighton Buzzard and Bragenham.

So I think it was 3 units gone. Because it was not in London were you have loads of people to pick, at Leighton Buzzard you only have a few BTP units out there. There was a BK unit on an address enquiry and the BG unit but he was single man, and there was a motorcycle units responding. Because of the numbers, I knew that the Milton Keynes unit was only one-man; I knew that he couldn't do it on his own. So I send the BK unit and the motorcycle, because the potential area for the search might be huge.

I: Did you wait for updates or did you prompt for more information? From whom? What did you do with this information?

O: I made enquiries with the Network Rail to get the driver so we can get a full account. Then, I got an update back. They couldn't find the body and it was getting dark. We needed body recovery units. I put in the AQ that we needed to contact the Thames Valley Police. The sergeant arrived to the scene. The local police took charge. It was an area right on the board of Bedforshire and Buckinghamshire. They told the NOC to turn off the tracks. They had an hour to recover the body and collect evidence from the scene.

We handed the command over to the local police. It turned out to be a suicidal incident. Later they found a suicidal note, in the car, his wife has left him.

I: Was there any stage in which you find difficult to process and integrate the information?

O: at the beginning is always difficult, one can only guess with very little information. The decision depends on the urgency of the voice on the other side.

TRANSCRIPT 12 - Supervisor

What information did you have available after the interruption?

I: When you go for a break, how do you know what information you need when resuming your task? Could you prioritise which piece of information is most important when resuming the task?

O: Say that I've been out for fifteen minutes or so, Jenny or whoever is my colleague would have been dealing with things in my absence. Any major things they have been dealing with, they will obviously brief me. So in the period that I've been away, there was a person under a train. That would be quite a major incident. They will be saying: there was a person that got caught under the train in Oxford Circus; we have this unit and that unit going. A little brief.

And secondly, the way I do it, to give me an idea of all the messages that have been going on: I will go through a quick search on my system, to see of all the messages that have been going on in the last hour; and depending on which section of the radio I have been dealing with, I would look at the incident that reflect on my area. So I look at my messages, so for example, I'm dealing with the underground, you see suicidal women that is quite serious, then H10 (hotel one zero) going. I will be aware of …

So I will be brief (a) by my colleague and (b) by the system

I: Are there any other lists where you get your information from?

O: Yes of course, the tagging list. Usually when resuming your task you go to the tagging list and check G1 incidents.

I: Can you search only by area?

O: I can, but depending on… Because I'm the radio supervisor, I have all the areas set up. It is my role today; I look at all the areas. But when you come in, you set up the area.

I: Who do you assign to be a relief operator?

O: The way the radio works, on this side one person fills in. Because the radio is quite intense, mainly if there is nothing serious, the radio dispatchers will have a basically break every hour for fifteen minutes. On this side of the radio there are 3 people covering one channel, so they will relief each other. So every hour they will get a break. And there, of course is me here who will assist any person or chasing up other calls. My role is really I make the radio traffic trade better, so for example, I have to keep an eye on my queues to make sure that the radios have been actually been put out timely, and have been put out correctly, also I have to make sure that appropriate resources have been sent to each job.

For example, this job at the main line at King's Cross [pointing at the NSPIS] which was a report of 10 people fighting and the London north operator sent 2 units to it, and I felt that because the amount of people over there I felt that you need further officers there. It should be my job to make sure that I grab other units from other areas. I would say to other area person, could you put a call out in your channel to support these 2 units at King's Cross.

So, my job is to make sure that appropriate resources go there at the right grade so we don't have people charging on in blue lights indeed if it is not an immediate call.

I: You have a specific number of resources per area, so how do you decide that you need resources from different areas?

O: If say for example we have a fight or a big disturbance, I would be able to use any resources I want to in the areas, because obviously it is a grade 1 call, incident ongoing.

But if there is a train crash, or a bomb explosion, something that is going to *drag* on long term, and obviously you will have to talk with the duty officers from the area and ask them the resources they have available, and how many we can spare, and they sort of create them into a seal…

We are going to need them for longer periods, like If it's a train crash, they are aware that obviously we need people there initially, but depending on how severe this incident is, we know that is going to go on for hours so we are going to need more resources aside.

So we are saying 2 things: I'm allowed to send them from all the 3 areas, that's not a problem. But obviously for long term use of resources, you have to speak to the duty officers from each area and collate how many officers we have on duty and hopefully over a period of time we will find out how many officers we are going to need and how long we are going to need them for. We might have to call people at home. But that would be my job, to go to each duty officer, and find what their resources are on duty, and obviously identify which ones are needed.

We are one force, but the reason they are assigned to one area is because they are targeting certain places, but obviously if there is a need for them to be at a particular place, you cannot leave the whole area unpoliced, which happens sometimes. The question is not getting people there; you have to stop people from going. Sometimes we have to cancel…

I: When you come back from a break, and have to cover the channel again, what do you look at? How long does it take to you to know what has happened?

O: The most you can be away is 15 to 30 minutes. I appreciate that a lot can happen, but generally speaking it doesn't take long to see 30 minutes of messages. As I said, when I'm taking over, the operator will brief me: I had a big fight in such and such.

The way I'll brief myself is that I will go into each of these messages and look at who is going where.

The idea I want to keep in my mind who is going where, in my head not just from the screen.

So I want to keep it in my mind which units are going for certain calls.

So if it comes back a cancellation, or if they got there and need more units, I want to know H92 calls me, I need to know that that unit was going to east Croydon, rather than search where they are in the system.

Generally speaking, what happens to a radio operator, you don't know what is going on at the start of the shift. Over the period of the day, you know where everybody is going.

I: When resuming a task, which incidents do your colleague briefs you on?

O: They will tell me only about ongoing incidents.

I: When you have a huge list, do you read each message?

I will flick through and read only G1 immediate calls

You see here the grade, IMM is for immediate, PR is prompt, PG police generated, NA non attendance, RO routine we have to deal with that in 24 hours, and DE is deferred. Say for example that we get a report that a person is on a train travelling from Manchester to Euston, but he is not going to arrive at Euston for two hours. He is causing some issue; he had stolen a can of beer from the bar.

There is no point to point in the dispatch queue now, and send officers now to wait at Euston, because the officer will be waiting for two hours. So we will defer it. So the reason we put the defer mark there, because there is no point of attending until

I: When it is deferred, and the time has come, how does the system notifies you?

O: We will do initially, we will put it out to whomever unit, so we will make them aware, and then we will schedule it in the system, we pin it, we have to make it manually. We normally put it out just to make people aware of what is coming, that there is something pending.

We can assign a unit, which means we can still send it to other jobs, or we can deploy it in the system, which means it is assigned to that one incident. Say for example, that a train company in Euston reported … that BK43 will deal with it, so we can assign it, but we can still use it for other jobs in Euston Station.

I: When you have high workload, and an interruption occurs, how do you recover from it?

O: Let me show you: so any messages that we use on the dispatch queue, we tag it to our queue.

If it is in the dispatch queue and it is ongoing you tag it. So there you go all the outstanding calls.

I: In this case, when you came back from your break, how do you know which changes have you missed?

O: I wouldn't say you miss changes. When I came back to the radio, I said to the officer "Hang on a second". So when I came back, I was dealing with London south area, I can see control area which messages are in my area. I can also see [pointing to the tagging list] which messages are still open, London south A LSA.

There might be a lot more calls that have been taken within this room, but these are the ones that we are dealing in the radio in this area.

We tag it to the queue so that means that I can have a quick look of what is going on in the area, and just click that button there which is the Tag List Queue and just see instantly which messages are going through our queue. All the other messages are in the system at the moment, I'm not interested…

Appendix XI. Incident Summaries

Leicester Square – Fatality: person under the train - April 2008

1. A radio call came from the LU Line Controller at 16:00 on a Friday about a person under a train in Leicester Square. At the same time active messages from the call takers.
2. The radio dispatcher graded that as a G1. The location is critical, it is close to peak time and it is Friday. He requested information about the platform number. He also informed the Network Operation Rail who are in charge of the underground operation.
3. He deployed one officer and six PCSO (Police Community Support Officers). He enquired from the LU manager about staff for crowd control and evacuation. The first officer on scene reported back with the first update at 16:08.
4. He created the incident and sent to the action queue a request for LAS, MET Police, LFB and social care.
5. The radio dispatcher requested CCTV.
6. The next update came at 16:12: an average flow of 3000 people currently in the station. Not sure how many in the platform. An update came in the tagging list that 800 passengers were in the train.
7. He sent more units – 10 officers were deployed.
8. He advised RAIB and CID.

Three Locks Bridge – Fatality: person under a train – 2009

1. A female officer at Euston called upset. I've heard about a person under a train at Leighton Buzzard. "Are you aware of anything?" The radio dispatcher tells her that he has not been advised of anything and requests information.
2. Passengers are asking about the train that has been cancelled and enquiring if partly is due with the person under the train at Leighton Buzzard. In the meantime, three units called at the same time to be assigned: "I'm on route one zero, I'm on route eta two zero" The radio dispatcher told the units to hold until the incident is confirmed.
3. He chose the BG [bravo golf] unit, but only on a G2.
4. He created the incident in NSPIS.
5. Then, he overheard a call taker speaking "Does anybody know where such and such bridge is?" He asked him what it was about, and the call taker informed him about a call from Network Rail Birmingham about a person under a train. With that information he confirmed the rumours and the incident location. He informed the duty officer.
6. Once the incident was confirmed, the radio dispatchers called the other two units: "Units now, now you can go". The incident location is "Three Locks Bridge", not Leighton Buzzard. It was between Leighton Buzzard and Bragenham. He sent a BK unit and a motorcycle because the potential area for the search might be huge. The radio dispatcher asked the call taker to contact the driver – details could be obtained from Network Rail.
7. The first unit reported. They couldn't find the body and it was getting dark. They needed body recovery units. The officer asked to contact Network Operations Centre (NOC) to turn off the current in the track.

8. The radio dispatcher put in the action queue a call to the Thames Valley Police. More units were needed but there were no more free resources from the BTP.
9. The radio dispatcher checked the tagging list - the sergeant from Thames Valley Police arrived, and NOC confirmed the current was off. Once the current if off, NOC gives them an hour to recover the body and collect evidence from the scene.
10. They handed the command over to the local police who continued the search. It turned out to be a suicidal incident. Later they found a suicidal note in the car.

Gainsborough Road, E11 – POLAC – 08/03/2009

1. At 12:10 on the 8th of March 2009, an officer called the radio dispatcher and said “POLAC”. POLAC stands for police accident. The transmission stopped right after receiving the message. The radio dispatcher informed the duty officer about the incident: He shouted over 159 POLAC. The duty officer then plugged into the radio, listened to the transmission and monitored the incident.
2. Based on the call ID, the radio dispatcher knew it was an underground vehicle sent to an incident in Canary Wharf to support road closures. Not knowing an exact location but only where the unit was and where it was going to, he called for “urgent assistance” around that area.
3. Four units were deployed to search the area. Then, the radio dispatcher called the officer involved in the accident several times until the officer that was driving the vehicle confirmed the location. Then his transmission stopped.
4. The radio dispatcher confirmed the location with the four units and shouted over the room “LAS” (London Ambulance Services) required Gainsborough Road, POLAC. The officer involved in the accident called again requesting for an ambulance and the Metropolitan Police Traffic Unit. The radio dispatcher put these requests into the action queue for the call takers to contact LAS and the MET. The request for ambulance was made at 12:15.
5. At 12:16, the first unit arrived. The report was “Multiple casualties, officer unconscious trapped in the car”. The officer in scene asked for a Helicopter Ambulance due to critical injuries. The radio dispatcher put the update in the action queue and informed the ambulance crew about the multiple casualties. The radio dispatcher also requested for the Fire Brigade.
6. Check CAD and NSPIS for updates. From the CAD system, they knew

that the Fire Brigade was already aware. The first ambulance arrived at 12:21. The helicopter was requested at 12:23. The MET Police was aware of the accident and could give them a landing point for the helicopter. Five ambulances were on scene. They report possible neck and back injuries but LAS cancelled the helicopter at 12:27. Ambulance Service had sufficient people there to deal with the ongoing problem, the injuries were not serious enough to warrant a helicopter.

7. The area duty officer was informed. The first office (LE5) on scene was a sergeant but he required assistance. They called the area duty officer 15 seconds after LE5 was deployed with information about next of kin. A special unit was almost at location so they went to help. The special response unit A80 arrived. They have a camera on board. They were taking photographs of the cars, the final position of the cars, and the direction from where the vehicles came from and damages of the vehicles.
8. LE5 called back with information on the vehicles registrations: 3 vehicles involved. The radio dispatcher put it in the action queue for a PNC. He prompted LE5 for update on injured of the other cars and hospital names. The update received was that the driver of one of the vehicles was treated of shock. LA95 was going with the injured officer. There was a member of the public in the pavement. CID (Criminal Investigation Department) was informed and sent officers to the hospital. Thirty five minutes after the call came in, the Met Traffic Unit arrived.

Island Gardens – Train Derailment - 04/04/2008

1. The first action message was received at 6:15 am. It was an emergency call form the DLR manager and reported a railway accident.
2. The radio dispatcher graded the incident as a G2 because the train derailed.
3. He deployed four units. The first unit deployed was L22 at 6:23.
4. CCTV footage was requested.
5. The first update was 06:40: "Nothing suspicious, nothing terrorist related or criminal damage or vandalism. Somebody from the railway has left a piece of engineering equipment on the track. Traction current was off."
6. The radio dispatcher deployed a duty sergeant and 5 officers.
7. He advised the RAIB (Rail Accident Investigation Branch) and Pway (Permanent Way - a railway term for the track engineering crew.)
8. L22 report back that there are 59 passengers, the driver refused to blow the breathalyser and the DLR manager and sergeant advised to shut the track down until midday.
9. No injuries were reported, but the incident grade was updated to a G1, probably because of the effects on the surrounding community.
10. End of involvement. Three officers escorted the passengers to Deptford. No passengers were left on the train. CCTV was viewed and confirmed that equipment was left on the track the night before. Pway also confirmed that the item on the track was out on permit. Incident is handed to RAIB.

London Bombings – Radio Relief 7/7/2005

1. The radio dispatcher was assigned to "Radio relief".
2. Channel 2 (London Underground) went down.
3. Takes a tea break and decides to sort the channel out later.
4. He logged into the main channel received loads of active messages. He could not prioritize by grade because all were urgent so he did it sequentially.
5. The radio dispatcher informed the duty officer. Then, he deployed 3 units.
6. Then he created the incident and tried to keep the airwaves clear but active messages kept coming. Active Message from Edgware road "Numerous walking wounded, lost of lives and limbs". At the same time active message from Russell square "Walking wounded". The officers were asking for ambulances. The radio dispatcher put that into the action queue.
7. Then informed the fire brigade (LFB) assuming it was a train wreck.
8. He kept checking the action queue and the tagging list for updates on the 3 incidents: Liverpool Street, Edgware Road and Russell Square.
9. The grade was updated to immediate for all three incidents.
10. More units were sent to the 3 locations and LAS was informed about multiple casualties.
11. Other authorities were notified: the MET, LAS.
12. Then someone shouted "Bomb Damage in Edgware Road".
13. He prompted the officer on scene for information.
14. After 20 minutes from the first active message, a major incident was declared. Once a major incident is declared, gold, silver and bronze commands are set in place. The Gold Command is usually taken by the chief officer in the MET Police. The radio dispatcher assessed the best access routes for emergency services. BTP officers were prompt to

establish a rendezvous point (RVP) and a Joint Emergency Services Control Centre (JESCC). The MET Police was in charge of setting the outer and the traffic cordon. The LFB was in charge of the inner cordon.

The duty officers were informed about CHALETS[14], and the command was handed over to the Gold Command.

[14] Casualties – casualties, approximate numbers of dead, injured and uninjured. Hazards – hazards present and potential. Access – best access routes for emergency vehicles. Location – the exact location of the incident. Emergency – emergency services and other agencies present and required. Type – type of incident and brief details of number of vehicles, buildings involved. Safety – all aspects of health and safety and risk assessment must be considered by all staff working at or close to the scene.

Liverpool Street – London Bombings - July 7 2005

1. At 8:51, a call was received about a power outage at at Liverpool Street Station. There first update was a PCSO at Liverpool station stating there was dust or smoke. I told him to take a look if he can. So he came back with an update: a member of the staff said they had a power failure.
2. I informed the duty officer about the power failure. He was already aware. Then, I was trying to establish what happened: At Liverpool Street Station you have many lines, was it one line either directions, or all lines affected? second call came in: person under the train Edgware Road Station
3. He dispatched 2 or 3 units.
4. The radio dispatcher received a report stating dust or smoke. Then he overheard someone shouting Bomb Damage.
5. The radio dispatcher ran back to the radio, called the unit in place and asked him to confirm bomb damage. The PCSO on scene informed him that something suspicious happened.
6. He deployed more units to Liverpool and Edgware Road Station.
7. Then the supervisor asked him to monitor the handing over of the G8 Summit to the Glasgow Response Unit. The BTP was in charge of the security of the G8 but the incidents happening in the London Underground were more important. They needed all resources available. While the radio dispatcher was talking with the Glasgow Response Unit, he was also logged in the London Underground action queue to keep an eye on the active messages. He was also asked to tell all bronze officers about a telephone conference.
8. As soon as he handed over the command of the G8, he returned to his dispatch queue.
9. He checked the tagging list. The first report said “multiple injured, limbs missing, people covered in dust”; second report “other emergency services were in route”.
10. The Network Operation Centre (NOC) was informed about the incidents.
11. A radio call from an officer came in, a bus exploded at Tavistock Square. The

radio dispatcher put a call out to all units. They didn't have any more units but a BTP vehicle was on scene. The second update reported that doctors from the British Medical association were providing immediate help.

Appendix XII. Decision Tables

Leicester Square – fatality person under a train - April 2008

CUES	SITUATION ASSESSMENTS	ACTIONS	WHY?	WHAT FOR?
Call from LU controller (ICCS)	Fatality Critical Location Critical Time (close to peak hour)	Deploy units (6 PCSO + 1 unit) Request CCTV	Need officers + staff for crowd control, interviews and evacuation	
Radio update (ICCS)	Average flow 3000 in the station, not known number of people on platform.	Deploy 10 more units		
Action message turned blue (action queue NSPIS)	about 800 passengers in the train	Call RAIB and social care Advise CID	Train related incident	Need support for witnesses and investigation

Three Locks Bridge – Fatality: person under a train - 2009

CUES	SITUATION ASSESSMENTS	ACTIONS	WHY?	WHAT FOR?
Call from 3 units	A possible incident close to Milton Keynes	Check Tagging list and action queue for messages that confirm incident. Send one bravo unit	Cannot send units if incident is not confirmed	To enquire
Overheard call taker talking with Network Rail Birmingham	Accident confirmed – location Three Locks at Leighton Buzzard	Check location in the map (Gazetteer). Send the other 2 units that called before.		To confirm location
Radio update (ICCS)	Body cannot be found yet. Getting dark. Need to contact TFL NOC (Network Operations Centre) – need track current off.	Contact Thames Valley Police	Need more resources to look for the body	
Update (Tagging list - blue message NSPIS)	Sergeant from Thames Valley Police arrived. Confirmed track current is off.	Hand command over to the local police.	BTP has no more resources available.	

Gainsborough Road – POLAC - 08/03/2009

CUES	SITUATION ASSESSMENTS	ACTIONS	WHY?	WHAT FOR?
Call from officer (ICCS)	Police officer in traffic accident	Shout across LAS	Possible injuries	Ambulances require in the scene.
Call sign (ICCS)		See call sign from ladder	Double check unit involved	
Map (Gazetteer) + previous unit location (NSPIS)	Determine location	Call officer to verify but no answer. (1) Check NSPIS to verify previous location. (2) Check map to determine possible routes. (3) Check CAD to determine if there is another call about incident.		To send units to the right location ASAP
Update by radio (ICCS)	Multiple casualties. Officer unconscious trapped in the car.	Call for more ambulances + MET Police + London Fire Brigade (LFB)	To attend injured	
Action message turned blue (action queue NSPIS)	Helicopter needed. Advise Sergeant duty officer	Call LAS Call MET Police Call LFB Call Sergeant duty officer		To book helicopter For authorization for landing point Need to cut off top of the police car Needed on scene

Action message turned blue (action queue NSPIS)	3 vehicles involved.	Update NSPIS from CAD messages		
	Injured had been taken to hospital	Call Sergeant duty officer		To inform hospital name

Island Gardens – train derailment - 04/04/2008

CUES	SITUATION ASSESSMENTS	ACTIONS	WHY?	WHAT FOR?
Action message (red message in the action queue) – Incident notified by DLR Manager	Train Derailment – passenger train	Send units Call LAS	Might be injuries	
Radio update - Call from officer on scene (ICCS)	No injuries; traction current is off. Equipment left on track.	Call duty sergeant, LU offices and Pway		To notify that railway station has to be closed until midday.
Update (blue message action queue NSPIS)	Driver does not want to blow in breathalyser.	Send more units	Verify that equipment is from LU	

London Bombings - 07/07/2005

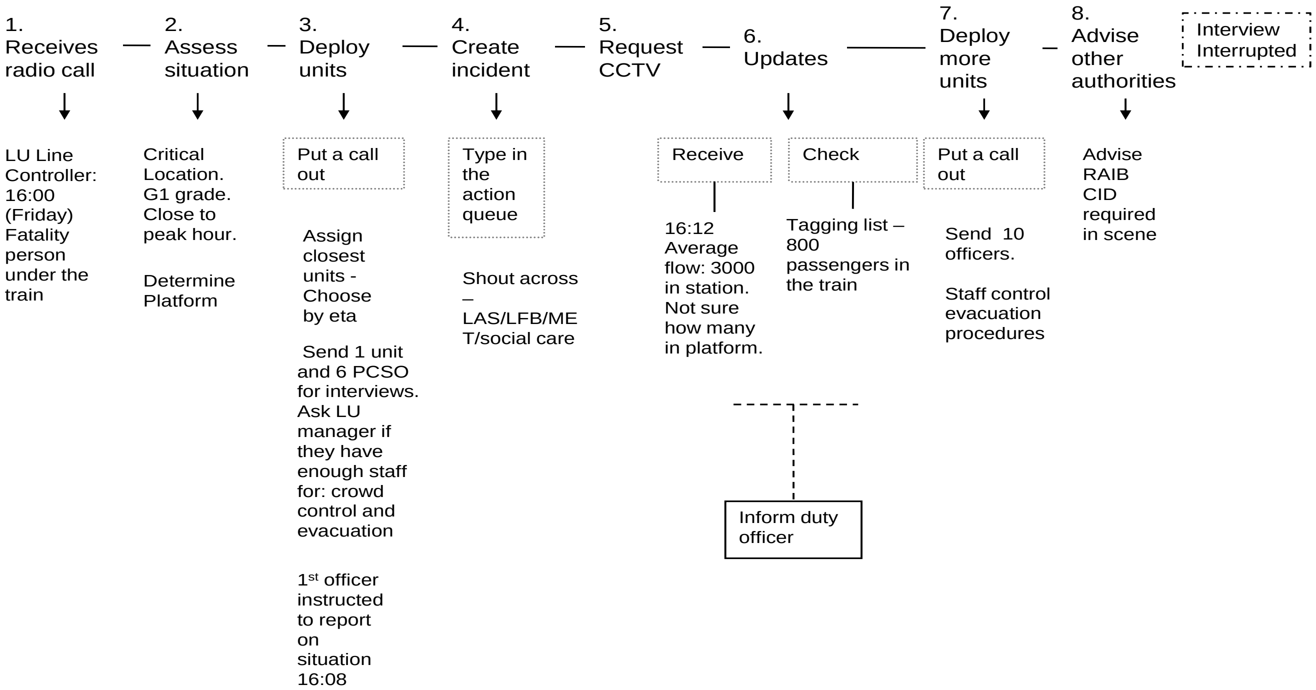
CUES	SITUATION ASSESSMENTS	ACTIONS	WHY?	WHAT FOR?
Channel 2 (LU) went down	Might be a major problem			
Action message (red message in the action queue)	Power outage Liverpool Street	Deploy 3 units Inform LFB Send LAS	An officer on scene is needed Multiple injuries	To get an on-scene report Medical assistance
Update tagging list	Reports from King's Cross. Leicester Square, Liverpool Street	Send more units Call duty sergeant, NOC, MET Police, RAIB, and CID	There are many injured, it is train related, and it looked suspicious	Provide assistance with crowd control. Assistance is also required for crime investigation.
Radio went down. Check CAD Updates	LAS and LFB on scene.	Major incident is declared	Four stations have reported incidents with multiple passengers injured	
Major incident declared "Bomb damage"		Prompt/Update RVP, JESCC and inner cordon. Determine best access routes for emergency services	Gold, Silver and Bronze Command has to be put in place	

Liverpool Street – London Bombings - 07/07/2005

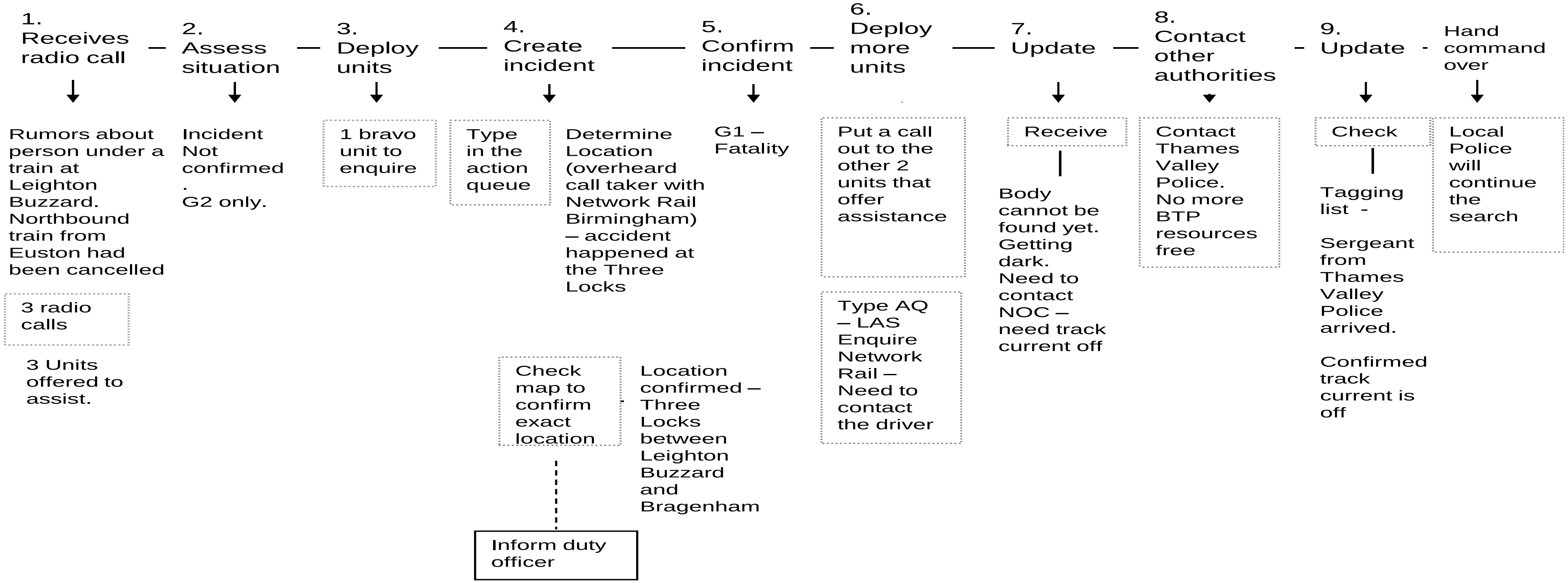
CUES	SITUATION ASSESSMENTS	ACTIONS	WHY?	WHAT FOR?
Action message (red message in the action queue)	Power outage Liverpool Street	Deploy 3 units	An officer on scene is needed	To report what is really happening on scene
Radio update - Call from officer on scene (ICCS)	Smoke or dust report Something suspicious	Call duty sergeant, LAS, MET Police		
Overheard another radio dispatcher	Bomb Damage Edgware Road	Deploy more units to both Liverpool and Edgware Station		To provide assistance
Other dispatch queues (when monitoring G8)			To keep situation awareness	To know what the situation is when finishing G8 hand over
Tagging list	Multiple Injured, limbs missing, people covered in dust	Inform NOC	Major disruptions in Central London	To manage the disruptions, technically and with the Press
Receives radio call	Bus exploded-Tavistock Sq.	No more available units.		To provide assistance

Appendix XIII. Incident Timelines

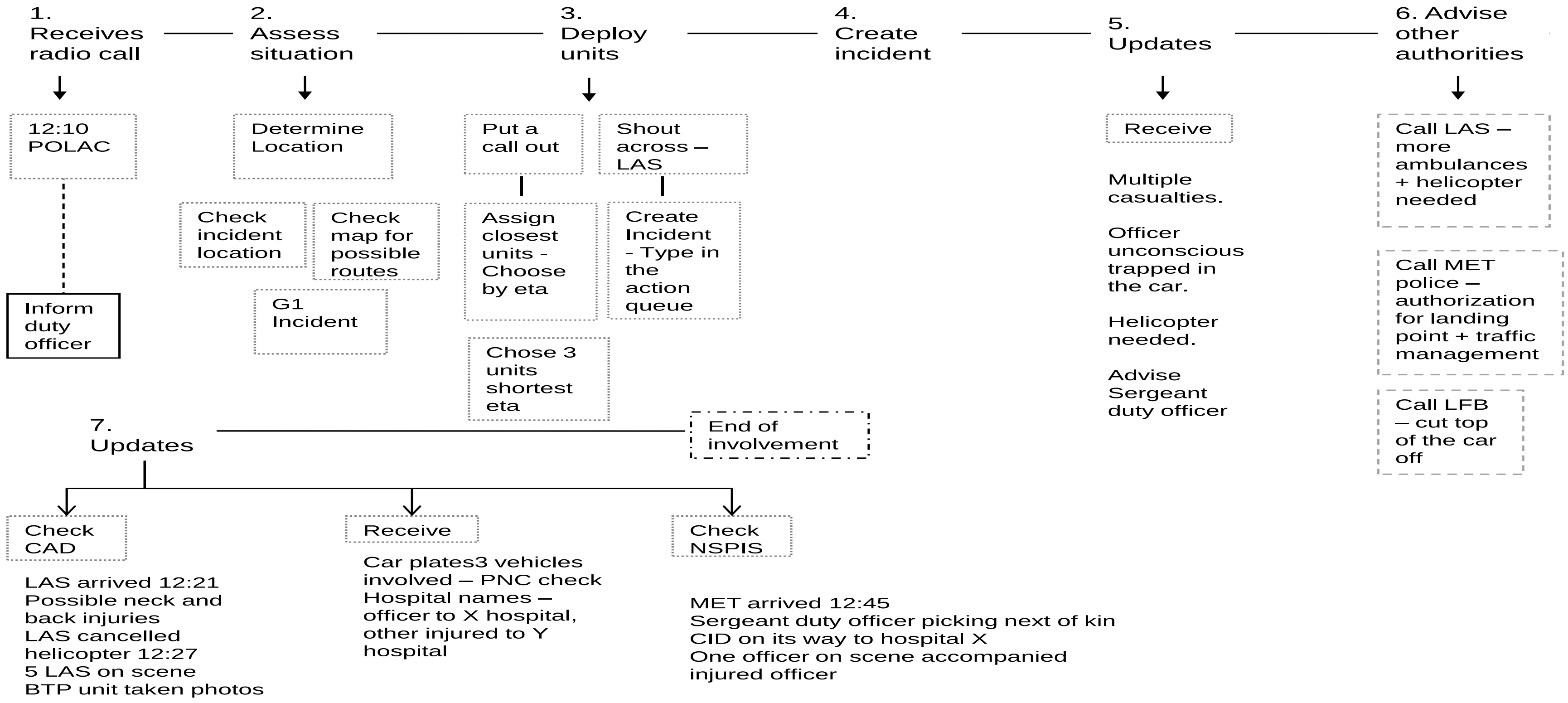
Leicester Square – fatality person under a train – April 2008



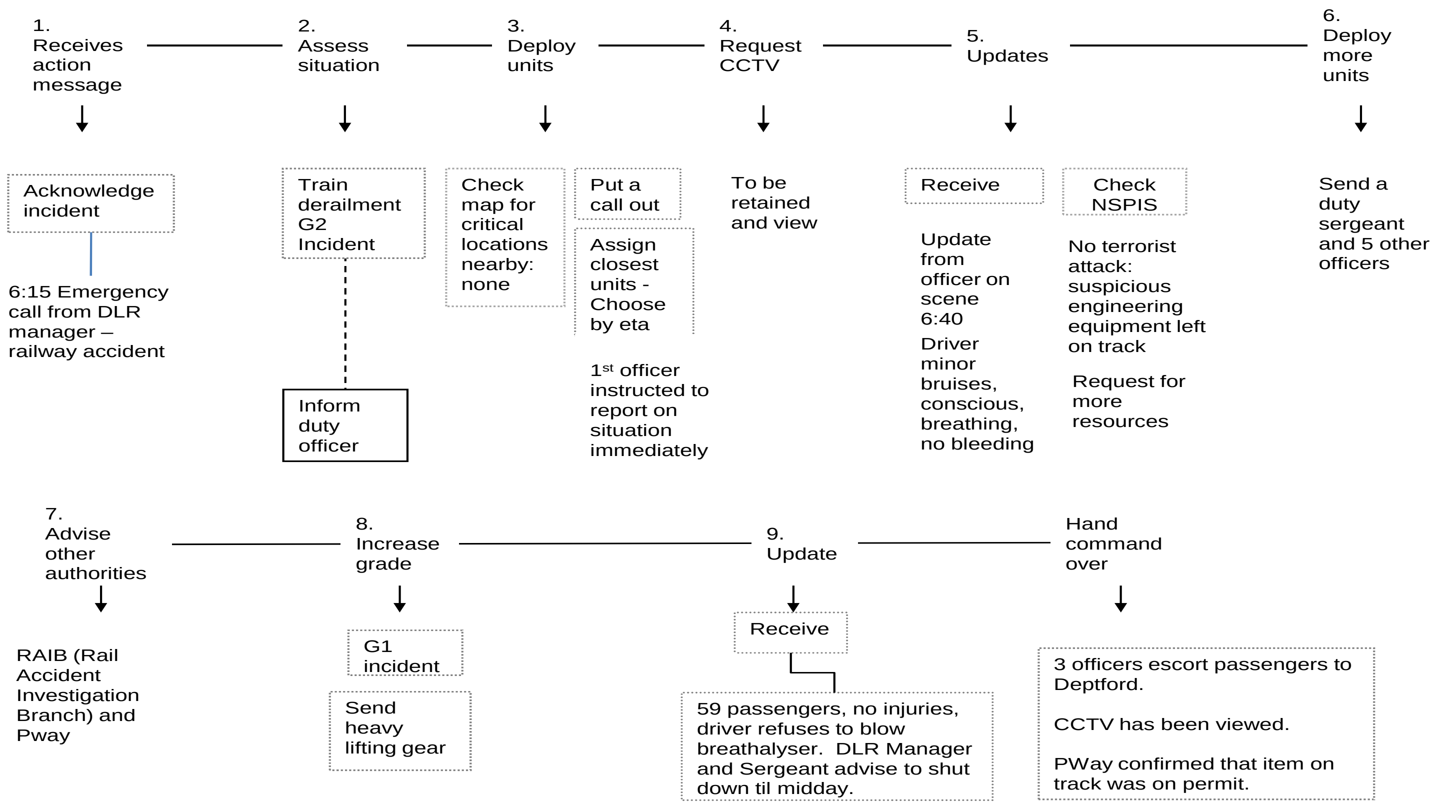
Three Locks Bridge – Fatality person under a train – 2009



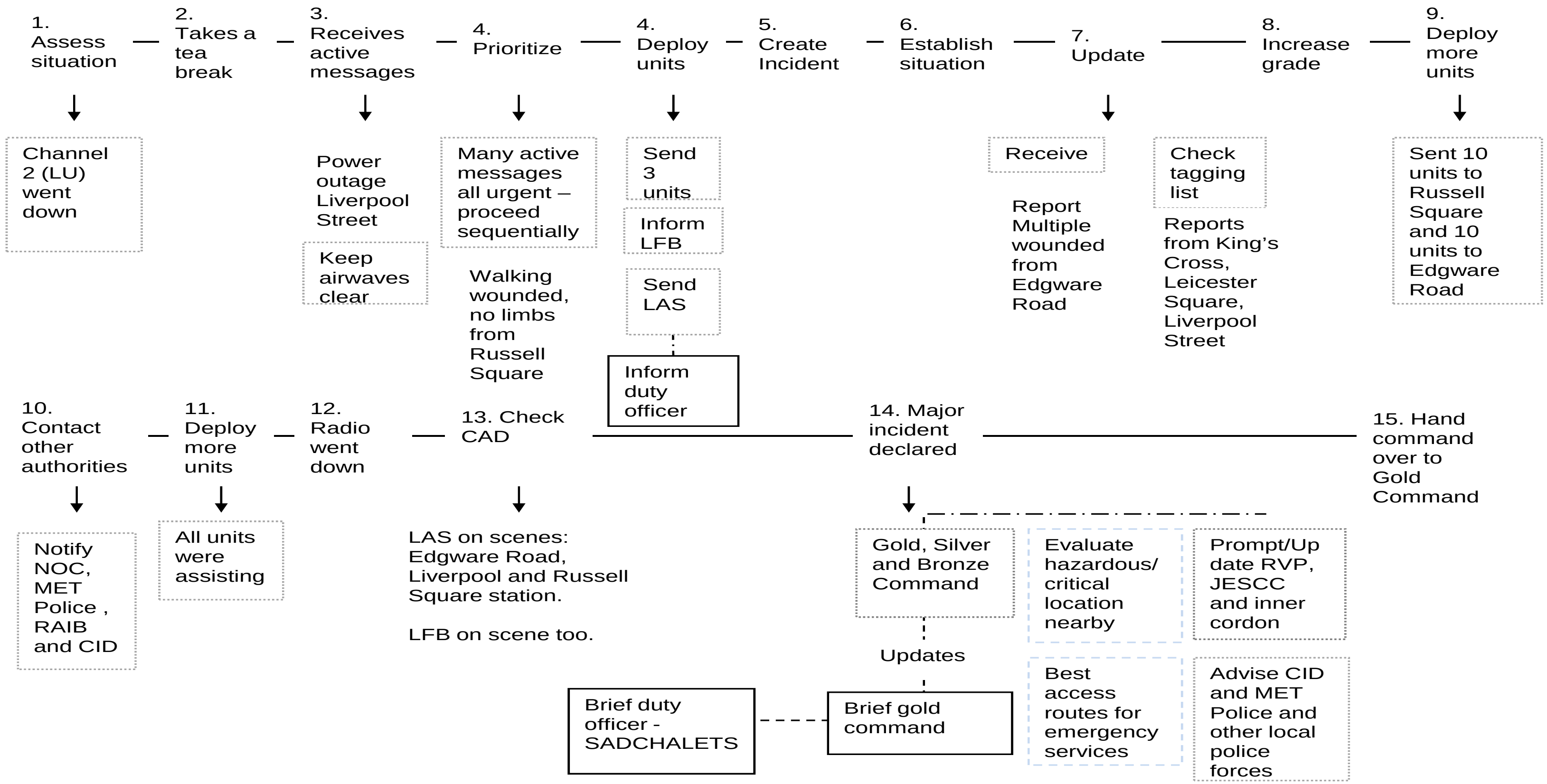
Gainsborough Road – POLAC – 08/03/2009



Island Gardens – Train derailment – 04/04/2008



Liverpool Street – London Bombings – 07/07/2005



London Bombings – 07/07/2005